1352

# UNIVERSITY OF IOWA
# STUDIES IN EDUCATION

CHARLES L. ROBBINS, Ph.D., Editor

VOLUME II NUMBER 7

# THE EFFECTS OF A SINGLE READING

## A Study of the Retention of Various Types of Material in the Content Subjects of The Elementary School after a Single Silent Reading

BY

GERALD ALAN YOAKAM Ph.D.

Edn. Research
G COLLEGE
11. 4. 52
No 1352

PUBLISHED BY THE UNIVERSITY, IOWA CITY

S.C.E.R.T., West Bengal
Library
Calcutta

S.C.E.R.T., West Bengal
Date 11.4.59
Acc. No 1352

# CONTENTS

*Chapter* *Page*

I THE PROBLEM . . . . . . . . . . . . 7

II THE NATURE OF SILENT READING . . . . . . . . . 9

III BACKGROUND FOR THE PRESENT STUDY . . . . . . 12
  Silent Reading Experiments . . . . . . . . . 13
  Summary . . . . . . . . . . . . 21

IV GENERAL DISCUSSION OF MATERIALS AND METHODS . . . . 22
  1. The Materials: Principles of Selection . . . . . 22
  2. Principles of Method . . . . . . . . . 23

V EXPERIMENTS WITH VARIOUS MATERIALS . . . . . . 25
  1. Materials . . . . . . . . . . . 25
  2. Method . . . . . . . . . . . 25
  3. Tests . . . . . . . . . . . 26
  4. Scoring . . . . . . . . . . . 27
  5. Experiments with the Article "Peanuts" . . . . . 27
  6. Experiments with Five Types of Reading Material Given to a Single Grade . . . . . . . . 37
    a. Civics Material . . . . . . . . . 38
    b. History Material . . . . . . . . . 39
    c. Narrative Material . . . . . . . . . 41
    d. Hygiene Material . . . . . . . . . 42
    e. Geograpical or Agricultural Material . . . . . 44
    f. Summary of Results . . . . . . . . . 44
    g. Further Conclusions from Experiments with Five Types of Material . . . . . . . . . 48
  7. General Conclusions for this Chapter . . . . . . 49

VI EXPERIMENTS IN WHICH VARIABLES ARE CONTROLLED . . . 50
  1. Purpose . . . . . . . . . . . 50
  2. Materials . . . . . . . . . . . 51
  3. Scoring . . . . . . . . . . . 52
  4. Method of Conducting Experiments . . . . . . 53
  5. Experiments with "The Admiralty Islanders" . . . . 57
    a. Introduction . . . . . . . . . . 57
    b. Interpretation of the Data without Regard to Grades . 57
    c. Establishing Corrections for the Variables . . . . 59
    d. Conclusions . . . . . . . . . . 60
    e. A Study of these Data when Distributed by Grades . 60
    f. Conclusions from Grade Distributions . . . . . 65
    g. Correcting for the Variables . . . . . . . 66
  6. Experiment with "Abbreviations" . . . . . . 69
  7. Conclusions from this Chapter . . . . . . . 71

BUREAU OF EDUL. & PSY. RESEARCH LIBRARY

VII THE EFFECT OF A SINGLE READING AND THE EFFECT OF TWO READINGS COMPARED . . . . . . . . . 74
  1. Purpose . . . . . . . . . . . 75
  2. Material and Method . . . . . . . . 75
  3. Results . . . . . . . . . . . 76
  4. General Conclusions Derived from this Chapter . . . 84

VIII RELIABILITY OF THE STUDY . . . . . . . . 85
  1. The Materials . . . . . . . . . 85
  2. The Method . . . . . . . . . . 90
  3. The Scoring . . . . . . . . . . 91

IX SUMMARY AND GENERAL CONCLUSIONS . . . . . . 94

X SELECTED BIBLIOGRAPHY . . . . . . . . . 99

# CHAPTER I

## THE PROBLEM

The purpose of the present study is to attempt to measure the effect of a single silent reading of various types of content material common to the elementary school. It is not a strictly technical problem, such as may be undertaken in the psychological laboratory, but rather an experimental study of a typical schoolroom activity in its natural setting.

It is commonly considered sufficient when a child comes to the recitation, if he can say he has "read" his lesson. Only the exceptional child can be depended upon to read or study his lesson more than once, or to use any other method of study than one or two readings of the material. Indeed, it is not by any means certain that it is desirable to read the lesson repeatedly until, as in the district school of an early day, it is "learned by heart." The relative efficiency of different methods of study has not been experimentally discovered. It is therefore important to determine the effect of a single reading of factual material, in order that, if this method of study be inefficient, other special methods may be devised to insure efficient work during the study period. Moreover this knowledge constitutes the first step in the study of the relative efficiency of various methods of teaching, such as the developmental method, the topical method, the problem method, and the socialized recitation, because these various methods depend for their success upon the basic step of one or more readings in the study period. And finally, the determination of the effect of a single presentation of material is fundamental in all experiments with silent reading, for without it the effect of various devices cannot be estimated nor the relation between the various factors in reading be clearly shown.

It may be pointed out also that a single reading is a typical procedure in life outside the school. When the business man reads his newspaper he ordinarily skims it or, at the most, reads some parts of it a single time. College students, pressed by social obligations and extensive assignments in History, English, Economics, and other subjects, are fortunate if they can read an assignment once. The busy pastor in his study must read rapidly a single time the many magazines and books bearing on his studies. The lawyer, the doctor, and the merchant find that they too have time for but a single reading. It is very important then to know how much can ordinarily be gained by this very common practice.

It is well known that there are many adults and perhaps some children who develop a high degree of skill in rapid reading and skimming. Ideally, it is the right of every person to have this skill highly developed. As a matter of fact we know very little about it. It is the purpose of this study, therefore, to investigate the effect of a single reading and to measure, if possible, the amount which may be reproduced by an elementary school child after having had a single contact with factual material when reading for information primarily and not for pleasure or appreciation.

# CHAPTER II

## THE NATURE OF SILENT READING

The most important investigations of silent reading are the contribution of the past two decades. For this reason we are still in the pioneer stage. Huey (1,p.5)* says, "To the average reader the process by which he gets his pages read is not very much better understood than was the performance of Livingstone by the savages. Indeed until twenty-five years ago science could not give a very much better specific account."

People have probably always read silently; but until the idea of *thought-getting* as opposed to *oral reading for the purpose of transmitting thought to others* came, no emphasis was placed on silent reading in the school. Students learned it incidentally through study. Even at the present time there is confusion among those who are working with silent reading as to its nature. This is quite clear to the student of scientific measurement in this field. So far, two factors in silent reading have been commonly measured, rate and comprehension, but it is only in the measuring of rate that any marked agreement is found.

Courtis (2, pp. 263-265) says, "In current practice the direct teaching of the mechanics of reading rarely extends beyond the third grade. Reading in the higher grades passes over into training in expression, in understanding, and in appreciation. Consequently ability in reading comes to have many meanings, each derived from the situation in which the term is applied. Ability in reading may mean:

1) Ability to recognize silently the general meaning of words of a given range of difficulty. (Otis)
2) Ability to sound correctly a given set of words. (Jones)
3) Ability to read aloud smoothly and with proper expression (without regard to whether the meaning is understood or not). (Gray)
4) Ability to read either silently or orally and to understand

---

* Refer to Bibliography for references thus indicated.

the essential relations existing between the essential elements of what is read. (Courtis)

5) Ability to read either silently or orally and tell in one's own words the substance of what has been read. (Starch, Brown, Gray)

6) Ability to read instructions either silently or orally and be able to act in accordance with the instruction. (Kelly)

7) Ability to read again and again (study) until one has mastered the contents of a passage so that one can answer questions about it or use the information in solving problems. (Thorndike)

8) Ability to read a passage and interpret the allusions it contains.

9) Ability to read a selection and be stirred emotionally by its aesthetic elements.

10) Ability to read a passage and interpret the mood, ideas, or ideals of its author.

11) Ability to read a selection and make judgments as to its style and merits as a piece of 'good English.'

And there are doubtless many other possible variations of the senses in which 'ability to read' may be understood.

Unfortunately, the makers of tests have not given much attention to this phase of the subject. They have labeled their productions 'Tests of Reading' and have been content to say explicitly or to imply that by reading is meant the ability to complete the test successfully. But when several such tests are given the same children, as at Gary, one seems to have the right to expect that the reading abilities of children as revealed by one test will show some agreement with abilities revealed by the next test, since both are tests of reading. This, however, is precisely what the results in general do not show.”

Reading includes many complex factors. The clearest distinction to be drawn between different kinds of reading is the universally recognized difference between oral and silent reading. Each in turn is further subdivided into types which are determined by the reader's purpose. There is oral reading for appreciation and information; generally the object is to give information or pleasure to others. Likewise silent reading may be roughly divided into reading for information or study, and reading for appreciation or pleasure. Each of these types may be again resolved into minor types. Silent reading for information includes skimming, rapid reading, and careful reading. Proof-reading is still another variety of silent reading, involving, however, merely the mechanical accuracy of the printed material. The purpose of the reader thus determines the type of reading which is employed.

In the present study the idea of reading as advanced by Courtis and Thorndike has been followed. Reading means the ability to gather thought from the printed page and to remember that thought long enough to answer specific questions about it, or to perform exercises which give evidence of an understanding of the facts contained in the material read. That is to say, careful silent reading is the type here under investigation.

It is recognized that silent reading is a very complex process. The factors in silent reading thus far recognized are rate, comprehension, organization, and remembrance. In this study the factors of rate and comprehension are involved particularly. Memory, both immediate and delayed, are prominently involved, since no measure of the effect of reading can be made without memory as a factor. As to whether or not the complex called reading is fully and adequately measured by these tests, or any others, this study does not attempt to state.

Tests of comprehension which make severe demands upon unstimulated recall are not favored by the writer. The most effective tests of comprehension so far seen are those of Courtis, Thorndike and McCall, Monroe, and Burgess. While not all of these measure the same thing, it is perhaps not desirable that they should, because each one probably measures some phase of comprehension in silent reading and each contributes something to the final solution of the problem of fully measuring silent reading ability.

VOCATIONAL & EDUCATIONAL GU
LIBRARY
No..................

# CHAPTER III

## BACKGROUND FOR THE PRESENT STUDY

The present interest in silent reading among supervisors, principals, and teachers is probably due to the stimulation from reading investigations of recent years. Much interest has been created through the work of Huey, Gray, Dearborn, Courtis, Monroe, and others, and an extensive literature on the subject is rapidly accumulating. The old style of article on reading which dealt with the literary content of the subject and the techniques of oral reading has within the past decade given way to a new type which is concerned chiefly with the various phases of silent reading. The new material has consisted largely of reports of researches in silent reading. One notable contribution is the *Twentieth Year Book, of the National Society for the Study of Education,* Part Two. These reports are generally of experiments which have been carefully conducted and in which one single variable factor in silent reading has been isolated and studied. The effects of various practices upon rate and the comprehension of reading material are reported. Recently, three notable method books (3, 4, 32) resulting largely from this research have been published. In view of the importance of these materials it is necessary to review them before embarking on an account of the present study.

There are three classes of experiments which have a bearing on the present problem: first, experiments with rote memory, some of which involve a single presentation; second, experiments with logical memory involving a single presentation or several presentations of the material; and third, experiments with reading material of some length made primarily in the standardization of reading tests or in the experimental study of some phase of silent reading. The first and second types of experiments have been made chiefly by investigators of the psychology of memory; the third has been largely the work of those interested in silent reading as a school art. Only the third type will be reviewed here.

### SILENT READING EXPERIMENTS

Experiments of the third type mentioned are much nearer the problem of this thesis than the first and second. They are essentially group tests of silent reading ability. Some of them were made for the purpose of measuring the effect of changing conditions upon the ability to read silently; others were conducted in connection with the standardization of reading tests. They have been carefully examined for any data which may have a bearing upon the present problem. The most important of these studies will be reviewed and a summary of the data made with such application to the present problem as may seem profitable.

Pintner (10) reports a study made to compare the relative efficiency of oral and silent reading as measured by comprehension of the material read. He gives the scores of 23 fourth grade pupils. They were allowed to read stories adapted from the *Arabian Nights* for two minutes and were then required to write down as much as they could remember. Each passage was analyzed and the main ideas were noted; the papers were scored for the percentage of the total ideas reproduced. The average percent of the total possible points reproduced in oral reading was 34; in silent reading 40. This experiment probably involved a single presentation of the material and it gives us an idea of what the probable average score for fourth grade pupils on this kind of material may be. It is to be noted that less than half of the total ideas in a narrative selection are reproduced when the pupils have read the article once only. This, then, is a rough measure of the effect of a single reading. There are, however, too few cases to make the data conclusive and the factor of previous knowledge of the material is not considered. It may also be pointed out that the type of reproduction favors the child who is skillful at composition.

Whipple and Courtis (11) report the results of an experiment with individual adult observers involving the skimming of material from Mangold's *Child Problems*. The data are significant. Six observers skimmed twelve passages. The average percentages of reproduction for the twelve selections were as follows:

| Observer | Range | Average |
|---|---|---|
| B | 50—80 | 59 |
| D | 30—80 | 53 |
| Fr. | 60—85 | 78 |

| | | |
|---|---|---|
| Gov. | 40—80 | 70 |
| Gst. | 50—70 | 58 |
| Sk. | 20—80 | 50 |

The above scores indicate that it is possible for adult college students or professors to attain skill in skimming material a single time which enables them to retain from 50 to 78 per cent on the average of the total ideas in a selection. A single careful reading of familiar material might result in much larger scores. What the ability to do this type of reading may be among school children remains to be investigated.

One of the earliest attempts to measure extensively the attainment of pupils in silent reading and the effect of training upon such ability is that of Waldo (12) in the Sycamore schools. The investigation included pupils in grades 3 to 7 inclusive. The material used for testing purpose in grades 5, 6 and 7, was the school paper, *Current Events;* in grades 3 and 4, the story of Daniel Boone in Eggleston's *Stories of Great Americans.* The pupils were given the material and were told to read for five minutes but not to attempt to memorize figures or to reread. At the end of the reading period, they were told to write all that they could remember; then they were further tested by ten questions covering the ideas in the material which were marked on the percentage basis. Tests of this kind were given in September and March and the growth in reading ability measured. The per cent of ideas reproduced at each period is given in the following table. It will be noted that the type of reproduction here is better controlled than that in the previous experiment. Ten questions controlled the reproduction to some extent and composition ability was not so necessary in this test as in the uncontrolled composition test. In Table 2 the scores representing the per cent reproduced are given.

TABLE 2

WALDO'S EXPERIMENT IN THE SYCAMORE SCHOOLS
PERCENTAGE OF REPRODUCTION AFTER
SINGLE READING

| Grade | September Per cent Reproduced | Mean Var. | March Per cent Reproduced | Mean Var. |
|---|---|---|---|---|
| 3 | 44.6 | 23.5 | 44.0 | 10.8 |
| 4 | 56.7 | 21.2 | 60.9 | 15.8 |
| 5 | 16.3 | 14.7 | 25.6 | 14.3 |
| 6 | 27.1 | 15.0 | 35.0 | 12.3 |
| 7 | 42.7 | 19.6 | 48.3 | 17.1 |
| 8 | 55.3 | 20.4 | 62.5 | 13.9 |

It will be noted that material used for testing the third and fourth grades was different from that used for the fifth, sixth, seventh, and eighth. The above scores do not therefore show growth from grades three to four. There was also no means used of equating the difficulty of the materials used for testing in the spring and the fall; nor does one know how much of this production is due to previous knowledge nor how fine a measure of the ability to reproduce the test represented. Nevertheless, one gets an idea of the power to reproduce after a single reading. This material will be used for comparative purposes later.

The experiments reported by Brown (13) throw some interesting light on the present problem. Brown gave reading tests to 400 pupils in seven school systems, including grades three and six. His tests consisted of a simple printed prose selection of a page and a half of material. Each child was directed to read for one minute and he was then asked to write all he could remember. This probably involved a single reading on the part of most children. The manner of scoring the papers was to mark the papers for the percentage of the total ideas correctly reproduced. The results show that the third grades were able to reproduce after a single reading on the average 30.57 per cent of the total ideas in the article; the sixth grades reproduced 27.37 per cent. Since the articles were not equally difficult, we have no means of measuring the growth of this power of reproduction after a single reading from the third grade to the sixth. Free reproduction of this kind also favors pupils with linguistic or composition ability. The data are, however, significant. Note that only one-third and one-fourth, respectively, of the selection are reproduced and these scores are uncorrected for the effect of previous knowledge. Brown does not give us any data for delayed recall in this experiment.

One of the most significant studies in reading made by experimenters within recent years is that by Gray (14) in the standardization of his oral and silent reading tests. The part most significant for the present study is that dealing with his silent reading tests. The experimentation which preceded the publication of these tests involved the testing of 3,161 pupils in grades two to eight inclusive. The materials used consisted of several selections of about 600 words in length, of which three selections were finally chosen for standardization, namely "Tiny Tad" for grades two, three, and four; "The Grasshoppers" for grades four, five, and six, and "Ancient Ships" for grades six, seven, and eight. These tests are given in such a way that a single reading is probably involved in most cases, for the pupils are told not to reread. The methods of testing were reproduction and answers to questions, which Gray says are justifiable because they are the two most commonly used methods of testing comprehension in our schools to-day. He also points out that a combination of the two methods is advantageous because some children do better by one method and some by the other. Oral reproductions are required below the fourth grade, because, as Gray points out, children in the second and third grades are greatly handicapped when tested by methods of written reproduction. The reproductions are scored by the word-counting method and the questions by giving a score of 10 per cent to each question answered correctly. In the reproduction method "all erroneous statements, all repetitions, and all irrelevant statements were checked from the pupils, reproductions and the remaining words counted. The ratio of these words to the total number of words read was taken as the reproduction score." In the second part of the test, a credit of ten points was given for each question answered correctly. "Theoretically this method of grading would be accurate provided the questions were of equal difficulty. Now, in fact, the questions are not equal in difficulty if we take the number of times each question was answered correctly as an index of difficulty. However, the relative difficulty of the questions differed for different schools. The experiences of individuals and classes differ widely, and it is legitimate to expect that some questions will be less difficult for some than for others." Gray finally hit upon the plan of determining the difficulty of the questions by examining the papers of the first 100 pupils to discover the number of times the questions were answered correctly. Questions of approximately equal difficulty were then chosen and given an arbitrary value of 10 points. The average of the two scores of each pupil, reproduction and answers to questions, was then taken as the comprehension score. This method has been criticized by some writers on the ground that the combined score does not tell definitely the relationship between what the child does and his maximum possible achievement on this test.

The data significant for the present study are found in Diagram IX, page 117 of the same monograph. This figure shows what the achievement of 2,564 pupils is after a single reading of the material "Tiny Tad," "The Grasshoppers," and "Ancient Ships"; but the writer knows of no way to interpret the data so that one may know what proportion of possible maximum achievement is here represented. The combined comprehension score needs for our purpose a reinterpretation which the data furnished in the monograph will not allow us to make. This is no doubt due to the fact that our particular problem was not under consideration. One generalization that we may reasonably make is that the scores presented here probably represent only a small part of the possible achievement which might be made if pupils were trained to be more efficient in comprehension of what they read when giving an article a single perusal.

A study was reported by Pintner and Gilliland (18) the purpose of which was to investigate the relative efficiency of silent and oral reading as a means of gaining ideas from the printed page. Their material consisted of short paragraphs of 50 words each, typewritten on cards. Each paragraph contained seven ideas. The cards were prepared in pairs of approximately equal difficulty, but on different subjects. One set of cards was used with high school students; one set in grades seven and eight; one in five and six; one in three and four; and a set of 16 cards was used with college students. One hundred and thirty persons in all were tested. The exercise was probably in the nature of a single reading since the observers were told to read the paragraph once only and then to write down all the ideas that could be remembered. The papers were graded by dividing the number of ideas reproduced by the time taken in seconds to read the article. Thus if five ideas were produced in 15 seconds, the score would be expressed as 33.3%. This method of scoring will bring out relative efficiency of the comprehension of pupils reading at various rates; but for our purpose the data given by these authors need a reinterpretation which is made in the following table.

TABLE 3

SHOWING PERCENTAGE OF REPRODUCTION AFTER A SINGLE READING

After Pintner and Gilliland

| | College | | High School | | 7 & 8 Grade | | 5 & 6 Grade | | 3 & 4 Grade | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 |
| Points | 5.02 | 4.58 | 3.98 | 4.68 | 5.66 | 5.47 | 4.7 | 4.76 | 4.52 | 3.71 |
| Percentage | 71.7 | 65.4 | 56.1 | 66.8 | 80.8 | 78.5 | 67.1 | 68. | 64.5 | 63. |

The percentage of the total ideas in these short paragraphs correctly reproduced after a single reading is shown to be high, ranging from 80.8% to 53.1%. This is no doubt due to the shortness of the article, because it is recognized that the longer the article the greater the difficulty of reproducing the ideas. The writer believes that the above data are significant for short paragraphs of material of the sort here used.

These same writers found that with a longer passage of 200 words containing 43 ideas the average number of points reproduced was 14.5 when reading silently, or 33.3% of the total. The per cent of reproduction is thus seen to be reduced either by the length or the difficulty of the passage.

Mead (19) reports an experiment the purpose of which was to discover whether pupils can reproduce more points after reading silently than after reading orally. Six tests of two minutes each were given to five different sixth grades, 112 pupils in all. The material was made up of stories from *Alice in Wonderland* contained in a 32-page pamphlet. The pupils were told to read the story for two minutes and the probability is that only a single reading was involved. They were then asked to write all that they could remember.

Scoring was on the basis of the per cent of total ideas reproduced. The results are summarized as follows:

| Class | Cases | Per cent reproduced | Mean variation |
|---|---|---|---|
| I | 20 | 25.22 | 6.6 |
| II | 24 | 48.6 | 11.9 |
| III | 20 | 29.9 | 7.5 |
| IV | 24 | 48.2 | 9.2 |
| V | 24 | 41.6 | 10.7 |
| Total | 112 | | |

The average for the five classes is computed as approximately 40.4%. This means that less than half the ideas in the article were reproduced after a single reading. The data are of course not corrected for the variable of previous knowledge and the exact nature of the materials read is not known, but the general findings are significant.

In a later report Mead also gives an account of an experiment similar to the above but involving some 340 children in grades three to ten inclusive. These scores are percentages of the total possible score after a single reading based upon number of ideas reproduced on a free reproduction test. The data are reproduced in Table 4.

TABLE 4

EFFECT OF A SINGLE READING AS MEASURED BY MEAD

| Grade | No. Pupils | Percent Reproduced |
|---|---|---|
| 3 | 20 | 34.2 |
| 4 | 20 | 23.9 |
| 5 | 18 | 45.5 |
| 5 | 21 | 43.1 |
| 5 | 21 | 35.7 |
| 6 | 20 | 29.7 |
| 6 | 16 | 44.3 |
| 6 | 20 | 45.1 |
| 7 | 20 | 45.1 |
| 7 | 20 | 37.3 |
| 7 | 23 | 38. |
| 8 | 20 | 66. |
| 8 | 20 | 56.9 |
| 8 | 20 | 58. |
| 8 | 20 | 55.1 |
| 8 | 21 | 6.7 |
| 10 | 20 | 56.6 |
| Total | 340 | |

It will be noted in the above table that there is considerable variation in the scores even with grades of equal rank, but that the general amounts reproduced range from slightly over half to less than one fourth of the total ideas in the selection. These data agree in general with those previously given.

Peters (20) reports a study which he undertook for the purpose of measuring the influence of speed drills upon the rate and effectiveness of silent reading. His experiment involved 207 pupils in grades three to six inclusive. Pupils in each grade were divided into two groups and measured in reading ability. One group then received drill in speed while the other did not. The tests to measure progress consisted of selections from 80 to 300 words in length, narrative in form, and easier than usual material read in the grade. The comprehension of the pupils was tested by written reproduction. The scores made on these tests are significant for the present investigation because the tests probably involved a single presentation of the material only. The quantity of reproduction 'is here reproduced in Table 5. These scores represent the total percentage of the complete ideas in the article which were reproduced by drill and non-drill groups.

TABLE 5

QUANTITY OF REPRODUCTION AFTER A SINGLE READING
After Peters

| Room | Nov. 6<br>% Reprod. | Jan. 6<br>% Reprod. | April 20<br>% Reprod. | June 2<br>% Reprod. |
|---|---|---|---|---|
| 3 wd* | 48.6 | 91.4 | 88.5 | 70. |
| 3 wd | 47.6 | 61. | 77. | 89.2 |
| 3 ad | 56. | 58. | 65.8 | 71. |
| 4 wd | 73. | 80. | 78.6 | 82.4 |
| 4 wn | 75.3 | 71.1 | 78. | 82.3 |
| 4 ad | 62. | 84.4 | 78. | 73. |
| 5 wd | 78. | 65.7 | 90. | 92. |
| 5 wn | 72.5 | 62. | 88. | 85.2 |
| 5 ad | 68. | 70. | 87.6 | 73.4 |
| 6 wd | 58. | 63. | 67.3 | 81.4 |
| 6 wn | 57. | 58. | 65. | 74. |
| 6 ad | 84. | 81. | 50. | 89. |

The above scores were made on narrative materials of 80 to 300 words in length, easier than material ordinarily read. The materials were not equated as to difficulty so we do not know the significance of these scores. They are of interest in the present study, however, because they give some idea of the probable amounts reproduced as the result of a single reading. This, of course, is only a by-product of Peters' study and does not represent a direct attack upon the problem.

In connection with her master's dissertation Mrs. Edith G. Germane (21) collected some data on a single reading. She examined 230 pupils using an article of factual character entitled "Peanuts". She measured comprehension by a test of short questions which were scored by giving one point to each correct answer. The data are summarized in the following table.

TABLE 6

SINGLE READING OF ARTICLE "PEANUTS"
After Germane

| Grade | No. Cases | Range | Av. %<br>Reproduced |
|---|---|---|---|
| 6 | 56 | 50-1.8 | 21.2 |
| 7 | 91 | 66-7.1 | 27.3 |
| 8 | 83 | 62-5.1 | 32.3 |
| Total | 230 | | |

The above averages were computed by averaging the data given in Mrs. Germane's distribution tables. It is significant to note that the percentage of reproduction after a single reading is small. These data are uncorrected for previous knowledge of material in the article and do not measure the effect of lapse of time. Mrs. Germane counted points as a measure of reproduction. These points have been changed by the writer into per cents of the total possible score.

---

* w and a represent different buildings; d, indicates a drill group; n indicates a non-drill group.

### 1. SUMMARY

The review of these studies reveals the fact that there are some scattered data which are significant in the solution of the problem at issue. These investigations are only indirectly applicable to the problem at hand. They deal with many different lengths of material, in different grades, employing different measures of reproduction, and different types of scoring. They do not completely answer our problem, "What is the effect of a single reading on the comprehension of factual material as measured by immediate and delayed recall?" for there are several variables involved, among which are the effect of previous knowledge of the material read, the different character of the material, its length, and the effect of these factors upon the power to reproduce, and the manner in which the testing is done. They deal, moreover with one type of material generally, namely, narrative material, whereas the present problem is to measure silent reading of factual material. No attempt was made in any of the studies to measure specifically the effect of a single presentation of materials of varied lengths. It will be the purpose of this study to contribute to our knowledge of these factors.

The data are summarized for the reader's convenience in Table 7.

TABLE 7

SUMMARY OF INVESTIGATIONS CONTAINING DATA ON EFFECT OF A SINGLE TEADING

| Investigator | No. Pupils | Grades III | IV | V | VI | VII | VIII | Character of Material |
|---|---|---|---|---|---|---|---|---|
| Pintner | 23 | | 40 | | | | | Stories Arabian Nights |
| Waldo | * | 44.6 | 60 | 25.6 | 35. | 48.3 | 62.5 | Current Events |
| Brown | 400 | 30.5 | | | 27.3 | | | Simple Prose |
| Pintner & Gilliland | 130 | 64. | 63 | 67 | 68 | 80 | 78 | 50 Word Paragraphs |
| Mead | 112 | | | | 40.4 | | | Stories |
| Mead | 340 | 34.2 | 23.9 | 43.1 | 44.3 | 38 | 55 | |
| Peters | 270 | 48.6 | 73 | 78 | 58 | | | Narrative |
| Germane | 230 | | | | 21.2 | 27.3 | 32.2 | Factual Article |

---

* Total number of pupils not given

S.C.E.R.T., West Bengal
Date 11.4.59
Acc. No. 1352

S.C.E.R.T., West Bengal
Library

# CHAPTER IV

## GENERAL DISCUSSION OF MATERIAL AND METHODS USED

The purpose of this chapter is to discuss briefly the materials and methods of the present study in order to give at the outset a general view of the problems involved. Details of particular experiments are left to the chapters in which these experiments are discussed. It is thought desirable here, however, to give the general principles upon which materials were selected, tests devised, and procedure determined.

### 1. The Materials: Principles of Selection

Since silent reading as involved in the present experiments is primarily a study of the printed page for the purposes of acquiring information, it is believed that experiments of this nature should involve the use of materials highly factual in character taken from the content subjects of the curriculum. Accordingly, materials were selected from various sources—geography, history, nature study, economics, and language. One narrative selection was used for comparative purposes. These materials were carefully arranged and adapted. In some cases materials already prepared by other experimenters in silent reading were used. Several selections and tests were prepared and arranged by the writer. Materials of different character were used in order that data of several types might be gathered. Acknowledgment is made in the later chapters of the use of borrowed materials.

Another principle of selection which is important is that materials of varied length should be used in experiments to determine the effect of a single reading. For this reason, articles varying in length from 100 to 2000 words were selected.

It is believed unnecessary to *hand-pick* materials to be used in these experiments because such materials as the pupils have put before them in the ordinary school text book are not always closely adapted to their understanding. The effort was made to select material representative of that ordinarily found and then to measure the ability of pupils to master the material at a single reading, regardless of any difficulties due to hard words or obscure meanings, which might be met with. It is also believed that grade progress would be more clearly shown by using the same material throughout.

In selecting the type of test to be used, the experimenter proceeded from the very first in the belief that a detailed test, objective in nature, requiring a minimum of composition on the part of the pupil, and objective and simple in method of scoring should be used. With this principle in mind for the main experiments, the controlled completion test and the true-false, or alternate chice test, were selected.

### 2. Principles of Method

Throughout the whole series of experiments the following principles were kept in mind:

1) School children vary greatly in ability to understand directions; therefore, great care should be taken to be simple and thorough in all statements as to what the experimenter wants done.

2) Children are often confused by the entrance of a stranger into the room and adjustment must be made before the tests are begun.

3) In attempting to measure the effect of a single reading, the variables of the situation within reasonable limits must be carefully analyzed and controlled.

4) When children take up a selection for reading they already know something about that selection. This is an important variable which ought to be measured.

5) The method of procedure in giving the experiments should be kept uniform throughout and variations from the method should be noted and recorded.

6) Written directions must be made simple, and oral instructions must be effective.

7) Practice with test material gives training in the use of such material; practice is a variable which affects the final scores and should be measured.

8) Previous experience with a test gives a certain set or attitude toward a reading of an article which that test measures,

and the effect of this set or attitude upon a single reading should be measured if possible.

In some of the experiments the factors mentioned under 5, 6, 7, and 8 varied somewhat throughout but where variations occurred they are noted and reported.

# CHAPTER V

## EXPERIMENTS WITH VARIOUS MATERIALS

This chapter will give an account of a series of experiments with various reading materials in the University Elementary School and in the grades of the North English (Iowa) Public Schools, involving 123 pupils in grades 4, 5, and 6. The pupils of the University Elementary School were tested on several different types of material so that the actual number of different experiments is quite large and the records will present the scores of 58 individuals in several experiments. Or to put it more plainly, 123 pupils were tested twice; 16 pupils tested three times; 14 pupils four times; and 16 pupils were tested three times on five experiments, making a total of 590 pupil test-units in all.

The purpose of this series of experiments was to gather data on the effect of a single reading upon several different types of material, but with the use of a comparatively rough but economical method of experimentation. The writer believed that data on several types of material would be more valuable for preliminary study, even though gathered in a somewhat rough way, than intensively worked out data on only one type of material. Accordingly, this series of experiments was undertaken as the first step in the process of measuring the effect of a single reading.

### 1. Materials

The materials selected for this series of experiments are of varying length, highly factual, and adapted to the purpose in hand. Some were printed, some typewritten, and some hectographed for use. In some cases, the articles were used as written by the author; in other cases they were adapted and rewritten by the experimenter. Each article will be described in the section of this chapter under which the scores in connection with it are reported. Some materials were prepared by other experimenters.

### 2. Method

The general method of giving the experiments in this series was as follows:

1) First a test was given to determine the amount of previous knowledge of the material the pupil possessed.

2) The pupil was given the article and told to read it *only once*.

3) At the same time the article was given out, the pupil was given a second copy of the test and was instructed to do the exercise (test) a second time just as soon as he had finished reading the material.

4) To insure that a single reading should take place, pupils were instructed to "Mark" by placing a line between the last word read at any given moment and the next word.

5) The general oral and written instructions (where written instructions were given) were the same as for the experiments in Chapter VI.

6) Pupils were not told that the second test (immediate recall) was to be the same as the first; but where the experiment was repeated, pupils soon came to suspect that the tests would be the same.

7) In order to keep up interest, the pupils' scores were reported to them after each experiment.

8) Pupils were told that a delayed recall would be required, but they did not have access to the article containing the material. Their delayed recall test was the same as the previous tests.

### 3. Tests

The tests used in this series of experiments were in the main devised by the writer. Where a test was adapted from some other source, acknowledgement is made. Two well known types of tests were used:

1) The controlled completion test in which the key word in the statement was suggested but was listed with four other words, so that it was necessary for the pupil to cross out the word, or words, in a statement which would make the true meaning clear. This type of test requires no writing. There is theoretically one chance in five for a pupil to guess the answer. The test as devised, however, is very rigid and includes practically every idea of any importance occurring in the original selection. This test was devised as a better measure of the

effect of a single reading than that type of test which leaves too much of unstimulated recall.

2) Controlled test of short statements or questions. This test requires much greater power of recall than the controlled completion test and it also requires that the pupils compose an answer to the question. It does not, however, represent as fine a measure of comprehension as the controlled completion test. A discussion of its validity will be found in a subsequent chapter.

These tests were finally adopted after considerable observation of the performances of individual pupils with various types of tests in a preliminary set of experiments which are not reported in this monograph.

### 4. Scoring

The method of scoring used in connection with the above tests is the same throughout: one point is given for each correct answer. The manner of determining the points was to follow closely the original article in each statement made and to devise a test statement or question for each important idea. Since the controlled completion test is much more detailed than the controlled question or statement test, the former involved more points and was more laborious to score. The manner of arriving at the final score was to find the total number of points in the test and to divide this number into the total number of points correctly answered. The final score was then expressed as a per cent of the whole number of points. Thus, if there were 100 points in the test and 60 were answered correctly, the score would be expressed as 60 per cent. This method of scoring has the advantage of being very objective. The relative difficulty of each point was not determined.

### 5. Experiments with the Article "Peanuts"

The material used to test the effect of a single reading in this series of experiments was prepared by Dr. Chas. E. Germane (23) for use in his problem, the effect of summarizing as opposed to re-readings on the comprehension of pupils in the public schools. The article is printed in pamphlet form on book paper and designed to duplicate the printed page. The booklet contains about 1900 words of highly organized factual material well adapted to the reading ability of pupils in the intermediate and grammar grades. The original article is found in Farmer's Bulletin No. 431, issued by the United States Department of Agriculture. Dr. Germane submitted this material to rigorous preliminary testing in which he eliminated all difficulties which might prevent comprehension of the article. It is, therefore, much more carefully prepared than the reading material ordinarily found in school texts.

The test used to determine the pupils' reaction to this type of material was devised by the writer. It is a controlled completion test of 25 statements involving 41 points which cover the main ideas in the selection. The pupils merely checked or crossed out the words which seemed to them to complete the true meaning. They were instructed not to guess at the answer.

The method used in testing is the same as that outlined in the first part of this chapter. No departures from the established method were recorded.

The tables which follow give the results of experiments with 156 pupils in grades 4, 5, and 6 of the University Elementary School and in grades 4, 5, 6, and 7 of the North English Public Schools.

TABLE 8

COMPARISON OF GRADE AVERAGES ON ARTICLE "PEANUTS" IN TWO SCHOOLS

| Grade | Before Reading | | Immediate Recall | | Delayed Recall |
|---|---|---|---|---|---|
| | *Sch. 1 | Sch. 2 | Sch. 1 | Sch. 2 | Sch. 1, 54 days |
| 6 | 30.8 | 18.7 | 62.1 | 37.2 | ------- |
| 5 | 21.2 | 13.3 | 50. | 27.6 | 39.3 |
| 6 | 12.8 | 12.4 | 37.6 | 18.4 | 21.5 |

The gain made as a result of a single reading of the article, "Peanuts", may be summed up from the above table as follows:

Grades 6

School 1—31.3% gain as a result of a single reading
School 2—18.5% ,, ,, ,, ,, ,, ,, ,, ,,

Grades 5

School 1—28.8% gain as a result of a single reading
School 2—14.3% ,, ,, ,, ,, ,, ,, ,, ,,

Grades 4

School 1—14.8% gain as a result of a single reading
School 2— 6. % ,, ,, ,, ,, ,, ,, ,, ,,

The above figures are the effect of a single reading as measured

---

* School 1—Pupils highly trained in silent reading
School 2—No silent reading

by immediate recall. Only *two* delayed recall tests were given in connection with this series of experiments, namely those in grades 4 and 5 of School 1. The results may be stated as follows:

Grade 5—Loss of 10 % as a result of a lapse of 54 days
Grade 4—Loss of 16.1% ” ” ” ” ” ” ” ” ”

Thus it appears from the average scores that the loss through lapse of practically two months is fairly small. However, it must be remembered that the children, before reading, knew considerable about this article. In grade four, after 54 days, the pupils knew only 8.7 per cent more than before reading; in grade five, only 18.1 per cent. Hence, the effect of reading a single time was not so great as might appear. Had 54 more days elapsed and had pupils forgotten at the same rate, very little effect from the single reading would have appeared. It must be remembered that the average score means that 50 per cent of the cases were below the grade average. Therefore, many children after 54 days knew no more than they did before reading. See Table 9.

The distribution of scores by individuals in the 5th grade, showing scores made on both immediate and delayed recall, will be given in Table 9 in order to show the individual variations around the mean.

### TABLE 9

INDIVIDUAL SCORES ON ARTICLE
“PEANUTS” IN GRADE V SCHOOL I
SHOWING INDIVIDUAL VARIATIONS FROM THE MEAN

| No. of Pupils | Rate | Before Read. | Im. Rec. | Delay. Rec. |
|---|---|---|---|---|
| 1 | 197 | 19.6 | 86.3 | 65.8 |
| 2 | 148 | 24.2 | 75.7 | 60.9 |
| 3 | 169 | 40.9 | 68.1 | 60.9 |
| 4 | 138 | 27.2 | 68.1 | 34.1 |
| 5 | 97 | 22.7 | 66.6 | 58.5 |
| 6 | ---- | 23.2 | 66.6 | 63.4 |
| 7 | 402 | 27.2 | 53.7 | 29.2 |
| 8 | 94 | 19.6 | 53.3 | 26.8 |
| 9 | 118 | 12.1 | 50.0 | 43.9 |
| 10 | 251 | 22.7 | 36.3 | ---- |
| 11 | 142 | 18.1 | 36.3 | 41.4 |
| 12 | 130 | 9.0 | 25.7 | 9.0 |
| 13 | 373 | 3.0 | 25.7 | 19.5 |
| 14 | 184 | 25.7 | 22.7 | 17.0 |
| 15 | 127 | 23.2 | 14.8 | 19.5 |
| Average | 170 | 21.2 | 50. | 39.3 |
| Average (14 pupils) | | | 45.6 | 39.6 |

The above table reveals a wide range in the ability of the pupils to reproduce the material in this article even after a lapse of 54 days (delayed recall). The table shows wide individual variations in each test. The range increased in the immediate recall test and decreased again in the delayed recall test. This is just what would naturally be expected, since the children are presumed to know but little of the material before reading; and the individual abilities of the pupils would operate to increase the range of their knowledge as a result of a single reading. Time would then operate to decrease the range again. Theoretically, there would come a time when the abilities of the pupils would be approximately the same as before reading. This does not mean that no impressions would remain upon the mind but rather that the ability to reproduce those impressions freely would have ceased. For a discussion of what happens to memory when individuals do not intend to remember materials permanently, the reader is referred to Meumann (6, pp. 313-335).

The above table would seem to indicate that the measurement of this ability is as successful in a general way as previous tests cited, since the usual phenomena connected with a silent reading test are found here: wide individual differences in ability to perform the reproduction exercise, and a somewhat symmetrical distribution of the scores about a mean. Moreover, the coefficients of correlation, as computed in a class in measurements, reveal the fact that the writer's tests correlate highly with both Monroe and Courtis.

It now seems desirable to combine the scores of the different grades as first given in Table 8 but this is not possible without first distributing them according to the frequency of occurrence or weighting the averages according to the number in each grade. The writer prefers to distribute the scores in the form of a frequency table which will bring out some results of the experiments in a clearer manner. Tables 10 and 11 show the frequencies of the scores in the combined grades of the two schools.

In making the combination of scores, the writer is aware that the abilities of the pupils in the two fourth grades, the two fifth grades, and the two sixth grades are not equal and that they may not represent the average sixth grade, fifth grade, or fourth grade. However, the danger is always present in a study of this sort and practical considerations move the investigator to make the combinations for the purpose of gaining whatever suggestions are possible.

It is also probable that in combining the scores into a distribution table where the class intervals are in steps of five that the resulting averages and means may differ from those in the original tables, but that the difference is so small as to be practically negligible. Moreover, the scores are not in themselves so fine that it is necessary to make exact distinctions.

Table 10 presents the scores of the pupils in the two schools as measured by the initial test given as a measure of previous knowledge. Table 11 presents the scores as measured by immediate recall.

### TABLE 10

DISTRIBUTION OF SCORES MADE ON TEST OF PREVIOUS KNOWLEDGE ON "PEANUTS"

Grades 4, 5 and 6

| | Two Schools | 123 Pupils | | |
|---|---|---|---|---|
| Interval | Grade VI | Grade V | Grade IV | |
| 50—54.9 | 2 | | | |
| 45—49.9 | | | | |
| 40—49.9 | 2 | 1 | | |
| 35—39.9 | 2 | | | |
| 30—34.9 | 1 | | | |
| 25—29.9 | 7 | 3 | 4 | |
| 20—24.9 | M_9 | 6 | 5 | |
| 15—19.9 | 7 | M_8 | 5 | |
| 10—14.9 | 7 | 9 | M_7 | |
| 5— 9.9 | 4 | 8 | 7 | |
| 0— 4.9 | 2 | 1 | 5 | |
| Total | 43 | 37 | 43 | 123. |
| Average | 21.9 | 15.9 | 10.5 | |
| Median | 20.8 | 15. | 9.8 | |
| Q | 7.4 | 5.4 | 5.8 | |

This table shows that in their ability to do the test, the pupils of these two schools are graded but that there is a great amount of overlapping and that the range of individual differences is wide. The scores are distributed in a symmetrical manner about the median. The P.E. in each case is rather small, showing that it is not necessary to go very far to each side of the median in order to include 50 per cent of the cases.

Since the pupils in doing this test were instructed not to guess at the answer, the averages may be taken as a rough measure of the amount which these children knew about the article "Peanuts" before reading took place.

In his subsequent calculations, the writer will use the foregoing averages to correct the medians for the effect of previous knowledge on the assumption that the pupils followed out instructions and did not guess.

TABLE 11

DISTRIBUTION OF REPRODUCTION SCORES A
AFTER A SINGLE READING OF "PEANUTS"

Grades 4, 5 and 6

Two Schools 123 Pupils

| Interval | Grade VI | Grade V | Grade IV |
|---|---|---|---|
| 85—89.9 | 1 | 1 | |
| 80—84.9 | 4 | | |
| 75—79.9 | 2 | 1 | |
| 70—74.9 | 1 | | 1 |
| 65—69.9 | 2 | 4 | |
| 60—64.9 | 1 | | 2 |
| 55—59.9 | 4 | 1 | 2 |
| 50—54.9 | 2 | 3 | 1 |
| 45—49.9 | M_7 | 2 | |
| 40—44.9 | 3 | 1 | 4 |
| 35—39.9 | 4 | 3 | 1 |
| 30—34.9 | 2 | M_4 | 1 |
| 25—29.9 | 3 | 5 | 4 |
| 20—24.9 | 2 | 3 | 3 |
| 15—19.9 | 2 | 3 | M_11 |
| 10—14.9 | 2 | 4 | 4 |
| 5— 9.9 | 0 | 6 | 6 |
| 0— 4.9 | 1 | 3 | 3 |
| Total | 43 | 37 | 43 |
| Average | 47.5 | 36.2 | 24.5 |
| Median | 46.2 | 31.8 | 18.9 |
| Q | 14.9 | 15.5 | 14. |

Table 11 shows that in their ability to perform this test after a single reading, the pupils vary widely. There is marked overlapping of the grades. Distinct grade progress is shown, however, by the fact that the median for each grade is higher than that of the preceding grade. The variation about the central tendency is about the same in each grade as is shown by the P.E. or Q which is 14.9, 15.5, and 14.

Tables 10 and 11 show in a very clear manner the effect of a single reading in at least three respects:

1. The average and median in each grade move upward after a single contact with the material.
2. The range is greatly increased.
3. The variability about the central tendency is increased.

That this change was due to the effect of a single reading and not wholly to a repetition of the test may be inferred from the following facts which resulted from an experiment to detemine the effect of a test upon a test. Twenty-five pupils in Grade 7 of the North English Schools were given the test on the article "Peanuts". Two and a half minutes were allowed to elapse and then the pupils repeated the test. The result was that on the initial test they made an average of 22 per cent of the total; and on the repeated test 21 per cent of the total possible answers. The pupils were instructed not to guess. It is possible that fatigue may have operated to cause this low second score but the time taken for the test was hardly long enough for this.

Before drawing the general conclusions which may result from a study of the foregoing tables, the writer desires to review at this point the scores gathered by Mrs. Edith G. Germane (21) in her study of the correlation between speed and comprehension in silent reading, to compare her results with those of the present study, and to sum up the conclusions which may be drawn from the combination of data.

Mrs. Germane measured the effect of a single reading on the article "Peanuts." Hers was a controlled test of short questions. It requires, however, more actual recall than the test used by the writer. The results of her tests are summed up in Table 12.

TABLE 12

GRADE AVERAGES ON ARTICLE "PEANUTS"

Germane's Study

| Grade | No. Cases | Range | Av. % Reproduced |
|---|---|---|---|
| 6 | 56 | 50%—1.8% | 21.2 |
| 7 | 91 | 66%—7.1% | 27.3 |
| 8 | 83 | 62.5—1.7% | 32.3 |
| Total | 230 | | |

The above averages are computed by averaging the data given in Mrs. Germane's distribution tables. The averages are computed by the same method as those used in the present study. They are not made as a result of the same test, however, and are not directly comparable with the averages given in Table 11 which sums up the results of the writer's experiments with the same article. However, the writer, after a close examination of the test made by Mrs. Germane, believes that the points covered are practically the same and that in a rough way the results may be compared to get a general idea of what the progress probably is from grade to grade in power to reproduce material of this type after a single reading. The chief difference in the two tests is that the scores of the pupils as measured by the writer's test are larger on the whole, and that the scores gathered by Mrs. Germane are uncorrected for the effect of previous knowledge.

The average scores of both experiments are brought together in the following table.

TABLE 13

COMBINED SCORES OF THE GERMANE AND YOAKAM EXPERIMENTS WITH THE ARTICLE "PEANUTS"

| Grade | No. Cases | Range | Germane | Yoakam | | |
|---|---|---|---|---|---|---|
| | | | | Init. | Immed. | Gain. |
| 4 | 43 | 0-70 | | 10.5 | 24.5 | 14. |
| 5 | 37 | 0-90 | | 15.9 | 36.2 | 20.3 |
| 6 | 43 | 0-75 | | 21.9 | 47.5 | 25.7 |
| 6 | 56 | 1.8–50 | 21.2 | | | |
| 7 | 91 | 7.1–66 | 27.3 | | | |
| 8 | 83 | 1.7–62.5 | 32.3 | | | |
| | 353 | | | | | |

The above table should be read as follows: In the 4th grade, Yoakam found that 43 pupils made scores which ranged from 0 per cent of the total ideas in the selection to 75 per cent; that the average score for this grade in the initial test, or test given before a single reading to determine previous knowledge, was 10.5 per cent; immediate recall after a single reading, 24.5 per cent, showing a gain of 14 per cent as a result of a single reading and the effect of the preliminary test.

The scores given above from the Germane experiments are immediate recall scores expressed as per cents of the total ideas in the selection.

The foregoing table is the result of the experiments given to 353 pupils in three school systems. The figures which may be in some degree comparable are those in columns 4 and 6. These scores are the results of a single reading on the same article but with different types of tests. It is possible to account for the higher scores in the Yoakam experiment by the fact that the test is easier, and the fact that the initial test was given which probably had some effect in motivating the reading and resulted in a somewhat higher score. The only check which the writer has upon the probable effect of the initial test upon the motivation of the reading is in connection with another series of experiments on a different type of material in which both the material and the tests employed by Mrs. Germane, were used, namely the material "Tuberculosis".

(See page 42). The result of the initial test for previous knowledge in Grade 6 (16 pupils was an average of 11 per cent with a range from 7.3 per cent to 50.3 per cent. With the same material and the same test in Grade 6, but with 55 pupils, Mrs. Germane recorded an average score of 36.5 per cent on immediate recall, or a slightly higher average than the scores gathered by the writer. These two sets of scores are:

Yoakam's Scores: Grade 6, Initial Test—11%; Immediate Recall—33.9%; Gain—23.2%.

Germane's Scores: Grade 6. Immediate Recall—36.5%.

The above scores do not show that the immediate recall is influenced to any great extent by the motivation resulting from the initial test; but the scores are suggestive of the probable amount by which the immediate recall in the Germane experiments is influenced by previous knowledge.

The writer is inclined to believe, however, that his own immediate recall scores are influenced to a considerable extent by the effect of the motivation of the reading due to the initial test of previous knowledge. It may be that, allowing for the easier type of test, for the effect of motivation due to the initial test, and for the operation of chance, the scores reported under the column "Gain" are not from what the effect of a single reading uninfluenced by any other variables might be. In that case, the figures reported by Mrs. Germane, combined with the gain as a result of a single reading recorded by the writer, may represent a rough measurement of a single reading in grades four to eight. In the absence of anything better, the following tentative measurements may be suggestive.

TENTATIVE GRADE AVERAGES AS THE RESULT OF A SINGLE READING OF THE ARTICLE "PEANUTS"

| Grade | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|
| Score | 14 | 20.3 | 25.7<br>21.2 | 27.3 | 32.3 |

The scores signify per cents of the total ideas in the selection. They are corrected in grades 4, 5 and 6 for previous knowledge and effect of motivation.

### *General Conclusions of Experiments with Material "Peanuts"*

The general conclusions which may legitimately be drawn from the experiments with the article "Peanuts" are suggestive but not conclusive. The reader is urged to make due allowance for the number of pupils examined and the uncontrolled variables which are no doubt affecting the totals. It is believed, however, that the results show more nearly the effect of a single reading than any other material available before this study was made.

The following are the main conclusions which may be drawn from a study of the tables presented.

1) The schools tested in these experiments differ greatly in reading ability as measured by the writer's test. (Table 8).

2) The gains made as the result of a single reading of the article "Peanuts" are from 6 per cent in grade four to 31.3 per cent in grade six. These scores represent actual increases in average scores made after reading the article once. The test was the same in both cases.

3) The gross average scores resulting from a single reading of this article, uncorrected for the effect of previous knowledge and for the effect of motivation in the reading upon the immediate recall scores, range from 18.4 per cent in grade four to 62.1 per cent in grade six.

4) The effect of previous knowledge or of guessing, as measured by the initial test, ranges from 12.4 per cent in grade four to 30.8 per cent in grade six. The writer is inclined to believe that this is largely the effect of previous knowledge or of defect in the test, although there may be a certain amount of guessing in these scores.

5) The lapse of 54 days between the immediate recall and the delayed recall tests given after a single reading of this article, caused a loss of 10 per cent in the fifth grade, and 16.1 per cent in the fourth grade. The gross scores were 21.5 per cent in grade four and 39.3 per cent in grade six. These scores are all the grade averages. (See Tables 10 and following).

6) The usual wide individual variations in ability to read appear in connection with these tests. The abilities are more closely grouped around the central tendency before than after reading, and the range is greater after a single reading than before. This is in accordance with what one would expect, since the better pupils would naturally make much larger scores after the reading and thus increase the range, provided, of course, the initial scores were low enough to permit such increases.

7) Overlapping is a marked characteristic of the individual scores in this series of experiments (Tables 10 and 11).

8) A comparative study of the scores gathered by Edith G. Germane with those of the writer, reveals marked differences, the latter's scores running much higher on material "Peanuts" than those collected by the former investigator. The writer attributes his higher scores to the type of test used and other factors.

9) Since in none of the experiments was an average of more than 40 percent recorded as a result of a single reading and this average, the result of at least two other factors than the reading itself, the experiments emphasize the need for some method of teaching other than the mere assignment of certain pages for the child's perusal during the study period. A single reading is not sufficient as a means of acquiring knowledge of this type of material.

10) Even if the assumption that a single reading is a common method of preparing for the recitation is not granted, still the figures resulting from this series of experiments are valuable as a starting point for further experimentation.

## 6. Experiments Involving Five Types of Reading Materials Given a Single Grade

In this section, the results of a series of experiments involving five types of materials given to Grade 6 of the University Elementary School are reported. The number of pupils examined in each grade varied from 15 to 17. The aim of this series of experiments was to get some idea of the relative difficulty of different materials as revealed in the individual and average scores and to form some idea of the relative efficiency of two types of tests, namely the completion test and the question-answer test.

The materials for reading were of a factual character and were gathered from different sources. They varied in length and character of subject matter. Some of the selections were printed in booklet form; others were mimeographed or hectographed. The individual selections are described as the scores are reported.

The method used in the experiments was the same as that reported in the previous section. Directions to pupils were uniform throughout. It is possible that in this series the pupils were trained in the method of the experiment to such an extent that they made better scores than they otherwise would have done. After the first test was given the pupils grew to expect that the test for immediate recall would be the same test as that which was given to test previous knowledge. This fact probably operated to raise their scores.

The tests used in this series of experiments were of two main types: a controlled completion test, and a test of short questions involving recall and a slight amount of composition. The results of the individual experiments are reported in the form of tables of the individual scores. It is thought that the averages in connection with so small a number of cases are not sufficient.

### *a. Experiment with Civics Material*

The reading material for this experiment was borrowed from Dr. Grover H. Alderman (25), then a graduate student in Education, who had prepared the selection in connection with an investigation as to the relative efficiency of the lecture method of the recitation and the question and answer method. The material was taken from a textbook and rewritten by Dr. Alderman. Together with the test, it was then mimeographed and made into booklet form. The type of test used is a controlled completion test in which the correct answer, together with three incorrect answers is included under a single question or statement. The test contains 20 points and theoretically there is an opportunity for chance to operate to the extent of 25 per cent, that is to say, the pupil has one chance in four of checking the correct answer in case he guesses. However, the writer does not believe that pupils guess in these experiments when instructed not to do so. The test involves no composition since the answer is indicated by checking the statement which the pupil believes to be correct. The length of the reading selection is 1200 words.

It is probable that the pupils knew very little about this material since it was about the government of Switzerland and was written for junior high school purposes. It contained more material than is generally found in the elementary geography.

Table 14 gives the rate and reproduction scores of 16 pupils in the sixth grade as measured by both immediate and delayed recall. The reader should keep in mind that in all the experiments of this section, there is an initial test to measure previous knowledge, a single reading, and an immediate recall. The delayed recall comes after a period varying from 36 to 57 days.

The class reported that the initial test in this experiment helped them with the immediate recall. They also reported that they had recently studied Switzerland but from a commercial, rather than a governmental point of view. The pupils were not told that they were to read the article following the initial test: neither did they know that the immediate recall was to be the same as the initial test.

TABLE 14

INDIVIDUAL SCORES MADE BY A SIXTH GRADE AS A RESULT OF A SINGLE READING "GOVERNMENT OF SWITZERLAND"

| No. of Pupils | Rate | Score Before Read. | Score After Reading | Gain | Delayed Recall. |
|---|---|---|---|---|---|
| 1 | 282 | 15 | 80 | 65 | 30 |
| 2 | 182 | 30 | 65 | 35 | 45 |
| 3 | 216 | 0 | 60 | 60 | ---- |
| 4 | 192 | 5 | 60 | 55 | --- |
| 5 | 504 | 25 | 50 | 25 | 45 |
| 6 | 186 | 0 | 50 | 50 | 25 |
| 7 | 182 | 0 | 50 | 50 | 30 |
| 8 | 204 | 5 | 45 | 40 | 25 |
| 9 | 186 | 0 | 45 | 45 | 10 |
| 10 | 184 | 10 | 45 | 35 | 10 |
| 11 | 444 | 10 | 40 | 30 | 35 |
| 12 | 196 | 10 | 35 | 25 | 15 |
| 13 | 142 | 10 | 30 | 20 | 10 |
| 14 | 546 | 10 | 30 | 20 | 15 |
| 15 | 214 | 5 | 25 | 20 | 25 |
| 16 | 174 | 10 | 25 | 15 | 30 |
| Av. | 253 | 9.8 | 44.7 | 36.8 | |
| Av.* | 259 | 10 | 43.9 | 33.9 | 25 |

### *b. Experiment with History Material*

The material used for this experiment is an adaptation of the description of a medieval castle, taken from Robinson and Breasted's, *Outlines of European History,* pages 387-392, and rewritten for the purpose by the experimenter. The article is 668 words in length, highly factual and interesting; therefore, it is particularly suitable for a silent reading experiment. The test used was devised by the writer, a partially controlled test of short questions or statements, with only a short answer required. There are 41 points in the test, divided among 29 statements or questions. Much more is required of the student in the way of power of recall and ability to compose his thoughts into short statements with this type of test than with any other used by the writer. There is very little opportunity for guessing, if any at all. In view of the more rigid nature of the test, as was to be expected, the scores are not as high as with the controlled completion test. The writer believes, however, that although the test is perhaps more typical of ordinary school tests than the controlled comple-

---

* The second average is for those who were present throughout the entire experiment: only these averages are used in succeeding tables.

Table 14 should be read as follows: On the initial test, the average score for the fourteen pupils in this grade was 10 per cent of the total possible score; on immediate recall, it amounted to 43.9 per cent of which 33.9 per cent represented a gain resulting from a single reading. After a lapse of 57 days a delayed recall was given with a resulting score of 25 per cent of the total possible answers, or a loss of 18.9 per cent. That is to say, these fourteen pupils remembered only 56.9 per cent as much after 57 days as they remembered immediately after a single reading.

TABLE 15

INDIVIDUAL SCORES OF SIXTH GRADE PUPILS AS A RESULT OF A SINGLE READING OF THE ARTICLE "MEDIEVAL CASTLES"

| No. of Pupils | Rate | Score Before Read. | Score After Reading | Gain | Delayed Recall. |
|---|---|---|---|---|---|
| 1 | 128 | 25.6 | 69.3 | 43.6 | 47.3 |
| 2 | 180 | 30.8 | 66.6 | 30.8 | 44.8 |
| 3 | 187 | 30.8 | 61.5 | 30.8 | 31.6 |
| 4 | 240 | 30.8 | 53.8 | 23.1 | 57.8 |
| 5 | 125 | 29.4 | 48.7 | 19.3 | 39.4 |
| 6 | 144 | 2.5 | 38.4 | 35.9 | 8.9 |
| 7 | 126 | 15.4 | 33.3 | 17.9 | 34.2 |
| 8* | 116 | 12.8 | 33.3 | 15.5 | -------- |
| 9 | 198 | 23.0 | 33.3 | 13.3 | 18.4 |
| 10 | 126 | 5.1 | 30.8 | 25.7 | 7.8 |
| 11 | 114 | 20.5 | 30.8 | 10.3 | 27.3 |
| 12 | 96 | 5.1 | 30.8 | 25.7 | 20.7 |
| 13 | 258 | 17.9 | 23.0 | 5.1 | 39.4 |
| 14 | 510 | 2.5 | 15.3 | 12.8 | 10.5 |
| 15 | 184 | 2.5 | 12.8 | 10.3 | 15.7 |
| 16 | 296 | 7.6 | 12.8 | 5.2 | 8.0 |
| Aver. | 251 | 16.3 | 37.4 | 20.6 | 27.3 |
| Median | 156 | | | | |

* This record is not included in the averages.

The above table should be read as follows:

On the initial test, the average score for the fifteen pupils in this grade was 16.3 per cent of the total possible score; on immediate recall, it amounted to 37.4 per cent of which 20.6 per cent represented the gain as a result of a single reading. After a lapse of 31 days, a delayed recall was given with a resulting score of, on the average, 27.3 per cent of the possible answers, or a loss of 10.1 per cent of the total ideas in the selection. In other words, these fifteen pupils remembered after 31 days 79.3 per cent on the average of what they remembered on immediate recall.

The reader's attention is called to the fact that some of the above pupils made better scores on delayed recall than on immediate recall. It is possible that reminiscence has been operating as a result of underlearning of this material.

tion test, it is not as good a measure of the effect of a single reading as the latter because it requires too much organizing power on the part of the pupil, and it demands more of the ability to recall. It does, however, rule out the guessing of answers. In measuring retention, it seemed desirable to use a test which would reveal associations of even a slight degree of potency which with the more difficult type of test remain unrecorded and unrecognized.

Table 15 gives the scores resulting from this experiment with 16 sixth grade pupils.

### c. *Experiment with Narrative Material*

The writer had a desire to find out what the results of a single reading of simple narrative material would be as contrasted with the results of a single reading of factual material. Accordingly, he selected from an old Harper's Fifth Reader a simple little story entitled, "Chasing a Rainbow." The length of the article was about 1000 words. A controlled completion test was prepared for the purpose of measuring the retention of the pupils of the same sixth grade which was examined in the previous experiments of this series. This test is of the same general nature as the other tests of the same type. The experiment was conducted in the same manner as the preceding experiments.

On the initial test the average score for the sixteen pupils in this grade was 34.8 per cent of the total possible score; on immediate recall, it amounted to 78 per cent, of which 44 per cent represented the gain as a result of a single reading. After a lapse of 39 days a delayed recall was given with a resulting average score of 70.3 per cent, or a loss of 7.5 per cent of the total ideas in the selection. To put it in other words, the pupils remembered after 39 days 90 per cent of what they had remembered on immediate recall.

The reader's attention is called to the fact that only one of the pupils in the grade made a better score on delayed recall than on immediate recall.

The writer accounts for the high scores made on the above material by the fact that the story was comparatively simple. It is also possible that the test had many potent or suggestive words in it that gave the pupil an easy key to the answer, and that the essential logic of the material led the pupil to make the correct choice in many cases. The high average score is significant that for these pupils, at least, simple narrative material is easier than material of a factual nature.

### *d. Experiment with Hygiene Material*

The reading material for this experiment was borrowed from that prepared by Dr. Chas. E. Germane (23) for use in his experiments. The article was adapted by Dr. Germane from a pamphlet, *What We Should Know About Tuberculosis* issued by the New York State Department of Health. The writer found the article very poorly organized and quite difficult because the facts were scattered throughout the article, which was 1900 words long, in such a way that it is doubtful if the average elementary or junior high school pupil could reorganize them so as to pass a very rigid test over the material. It is important that material for use with elementary school children should be well written. The poor quality of this article is not attributed to Dr. Germane but to the original writer.

The test used in this experiment was also borrowed from Dr.

TABLE 16

INDIVIDUAL SCORES OF SIXTH GRADE PUPILS AFTER A SINGLE READING OF HYGIENE MATERIAL "TUBERCULOSIS"

| No. of Pupils | Rate | Before Read. | After Read. | Gain | Delayed (36 days) Recall. |
|---|---|---|---|---|---|
| 1 | 366 | 18.8 | 50.3 | 31.5 | 25. |
| 2 | 210 | 15.6 | 46.8 | 31.2 | 37.3 |
| 3 | 186 | 6.2 | 40.6 | 34.4 | 15.6 |
| 4 | 306 | 19.9 | 39.3 | 20.5 | 40.6 |
| 5 | 150 | 12.5 | 37.5 | 25.0 | 28.1 |
| 6 | 204 | 9.3 | 37.5 | 28.2 | 25.0 |
| 7 | 174 | 0.0 | 37.5 | 37.5 | 5.0 |
| 8 | 198 | 28.1 | 37.5 | 9.4 | 34.3 |
| 9 | 210 | 15.6 | 37.5 | 21.8 | 18.7 |
| 10 | 366 | 18.8 | 34.3 | 15.5 | 21.8 |
| 11 | 144 | 12.5 | 28.1 | 15.6 | 21.8 |
| 12 | 288 | 6.2 | 21.8 | 15.6 | 12.4 |
| 13 | 138 | 6.2 | 21.8 | 15.6 | 15.6 |
| 14 | 660 | 6.2 | 15.6 | 9.4 | 9.4 |
| 15 | 198 | 10.5 | 12.5 | 12.5 | 9.4 |
| 16 | 168 | 3.1 | 7.3 | 4.2 | 18.7 |
| Aver. | 248 | 14.8 | 31.6 | 16.1 | 21.1 |
| Med. | 203 | | | | |

The above table should be read as follows:

On the initial test, the average score of the sixteen pupils in this grade was 14.8 per cent of the total possible score; on immediate recall it amounted to 31.6 per cent of the total possible score of which 16.8 per cent represented the gain as a result of a single reading. After a lapse of 36 days on delayed recall, the average score for the grade was 21.1 per cent of the total ideas in the selection. In other words, the pupils remembered after 36 days 67 per cent of what they remembered as measured by immediate recall.

Germane. It was also used by Mrs. Germane in her silent reading experiments. It is a very difficult test of short questions or statements calling for much memory work on the part of the pupil and for a reorganization of thought. Composition is also involved in some statements where, for instance, the pupil is told to "explain how tuberculosis kills a person." It is probable that with another type of test much higher scores would have been registered as a result of a single reading of this article. It is thought worth while to introduce the material here, however, because of the interesting comparisons which may be made and the resulting suggestions as to the possible effect of a single reading as measured by this type of test and with rather difficult material.

Table 16 gives the results of an experiment with 16 sixth grade pupils. The method is the same as that of the preceding experiments in this section.

The comparatively low score above may be due to the difficulty of the material, the difficulty of the test or to any one or more of a number of factors. The writer is inclined to ascribe it to the difficulty of the material and the difficulty of the test.

Table 17 gives in summary form the results of an experiment by Mrs. Germane with this same material. The only difference in method is that she did not measure the previous knowledge of the children. Theoretically, the children would know a great deal about this type of material because of the numerous campaigns of education which have been undertaken throughout the country from time to time.

TABLE 17

SUMMARY OF THE SCORES COLLECTED AS THE RESULT OF A SINGLE READING OF THE ARTICLE "TUBERCULOSIS"

From the Germane Study

| Grade | No. Cases | Range | Immed. Recall |
|---|---|---|---|
| 6 | 55 | 78.4—3 % | 35.4 |
| 7 | 94 | 66.6—13.6% | 39.3 |
| 8 | 78 | 69.9—18.1% | 45.3 |

The above table was made by computing the averages from the distribution tables in Mrs. Germane's thesis, pages 46-53.

For purposes of comparison, a summary of the scores collected by the writer follows:

TABLE 18

SUMMARY OF SIXTH GRADE SCORES RESULTING FROM SINGLE READING OF THE ARTICLE "TUBERCULOSIS"

Yoakam Study

| Grade | No. Cases | Range | Immed. Recall |
|---|---|---|---|
| 6 | 16 | 50.3—7.3% | 31.6 |

It will be seen that in the grades measured, the two experiments agree fairly well in the average score recorded. When it is remembered that the sixth grade examined by the writer performed correctly 14.8 per cent of the exercises on the average before reading the article, it is seen that the average scores recorded by Mrs. Germane are the result of factors other than the single reading, among which are the effect of previous knowledge, differences between the abilities of individuals of the two groups, the effect of the personality of the experimenter, differences in instruction, and other factors. In spite of these unmeasured variables, however, the results of the testing of these 243 pupils are very suggestive as to the probable effect of a single reading of the article, "Tuberculosis".

### e. *Experiment with Geographical or Agricultural Material*

The scores of the sixth grade concerned in the present study of five types of materials, were reported under the heading, "Experiments with the Article 'Peanuts'," in the first section of this chapter. They are brought forward at this time to this section for the purpose of comparison with the scores resulting from the other four selections given to this grade.

### f. *Summary of the Results of This Series of Experiments*

In order that the reader may have more clearly in mind the data resulting from this series of experiments, the following summary tables are given before any attempt is made to draw conclusions. The reader will note the varying character of the materials, varying in subject matter, in length, in manner of testing, and presumably in difficulty.

TABLE 19

SHOWING CHARACTER AND LENGTH OF ARTICLES USED IN SERIES OF FIVE EXPERIMENTS WITH A SIXTH GRADE

| Title of Article* | Character | No. of words |
|---|---|---|
| 1. Government of Switzerland | Civics | 1200 |
| 2. Medieval Castles | History | 687 |
| 3. Chasing a Rainbow | Narrative | 1000 |
| 4. Tuberculosis | Hygiene | 1900 |
| 5. Peanuts | Geography or Agriculture | 1900 |

The following table summarizes facts which should be kept in mind about the tests used to measure the effect of a single reading of the selections:

---

* The experiments were given in the order in which the selections are set down in the above table.

TABLE 20

TYPES ANE LENGTH OF THE TESTS USED

| | Name of Article | Type of Test | No. of points |
|---|---|---|---|
| 1. | Government of Switzerland | Controlled Completion | 20 |
| 2. | Medieval Castles | Short statements or questions | 41 |
| 3. | Chasing a Rainbow | Controlled Completion | 45 |
| 4. | Tuberculosis | Short statements or questions | 32 |
| 5. | Peanuts | Controlled Completion | 41 |

The foregoing table shows that in three of the experiments the same type of test was used and that in two others a different type of test was used. It is, therefore, impossible to compare the gross scores directly as they vary in amounts covered.

TABLE 21

SUMMARY OF DELAYED RECALL TESTS

| | Title of Article | No. days Elapsing | No. Pupils | Score | % of Immed. Score |
|---|---|---|---|---|---|
| 1 | Government of Switzerland | 54 | 14 | 25. | 56.4 |
| 2 | Medieval Castles | 31 | 15 | 27.3 | 79.3 |
| 3 | Chasing a Rainbow | 39 | 16 | 70.3 | 90. |
| 4 | Tuberculosis | 36 | 11 | 21.2 | 67. |

The above table shows that on the various types of materials the amount forgotten as a result of the lapse of time varies according to the kind of material and the length of time. It is thought significant that the apparent losses are not so great as one might be led to conclude they would be from the statement of Thorndike in regard to the desirability of over-learning anything which one would keep in mind for some time. There is an interesting problem in running down the amount of loss of various types of material due to the lapse of different lengths of time. The above table is suggestive that there is not so much loss on under-learned logical material as one might think.

Keeping in mind the limitations of the experiments as set forth in the foregoing tables, the reader will now interpret Table 22, which is a summary of the scores made in the various tests. It should be pointed out that there are a number of factors present here which probably affect the scores; namely, relative difficulty of the selections, length of the selections, type of test used, and differences in the length of the tests. The constant factors are the method of giving the experiment, objectivity of the scoring, and the effect of a single reading.

TABLE 22

SUMMARY OF THE SCORES MADE ON FIVE TYPES OF MATERIAL BY A SIXTH GRADE FOLLOWING A SILENT READING

| Name of selection | Initial | Immed. Recall. | Gain | Delayed Recall. |
|---|---|---|---|---|
| 1 Government of Switzerland | 10. | 43.9 | 33.9 | 25. |
| 2 Medieval Castles | 16.3 | 37.4 | 20.6 | 27.3 |
| 3 Chasing a Rainbow | 34.8 | 78. | 43.8 | 70.3 |
| 4 Tuberculosis | 14.8 | 31.6 | 16.8 | 21.1 |
| 5 Peanuts | 21.9 | 47.5 | 25.6 | ------ |

The above scores, as all the scores in this study, are expressed in per cents of the total possible score.

Confronted by the many differences which have been pointed out in regard to the above materials, the reader may be doubtful in regard to the relative difficulty of the different selections, or in regard to the relative amount retained as a result of the reading of the selections. Experience with the materials convinces the writer that such differences are clearly shown in the scores. A common method of showing such relations is that of correlation. If pupils are found to perform in such a manner that their scores in one test correlate highly with those in another, the two tests are believed to measure the same ability and the two materials are supposed to be relatively of the same difficulty. This opinion is expressed by L.W. and S.L. Pressey (26) in a study reporting the correlations on four different tests. They constructed "four scales each composed of a definite type of reading material. The items of three of them were taken from the Monroe Reading Scales and the 'Illinois Examination'; some of the items in the fourth 'scale' also came from these two scales. Each 'scale' had twelve items. One 'scale' was made up entirely of poetry, one of scientific passages, and two of ordinary reading matter, stories, and the like." These scales were arranged so as to be roughly equivalent and were then given to a large seventh grade consisting of 112 pupils. The poetical scale was given first, then the first general scale, then the scientific, and finally the second general scale. "By this arrangement, practice with the tests, fatigue and so on were so distributed that they could not act as a constant error in the results." The correlations were then computed and run as follows:

| | |
|---|---|
| Correlation of general 1 with general 2 | .85 |
| Correlation of general 1 and poetry | .38 |
| Correlation of general 1 and scientific | .35 |
| Correlation of general 2 and poetry | .31 |
| Correlation of general 2 and scientific | .49 |

Correlation of scientific and poetry .56

These writers conclude that the above correlations point out striking differences in the type of reading matter involved here and (by implication) in difficulty of these types. They say, "The exceedingly high correlations between the two general scales is first to be noted. The coefficient gives a measure of the 'reliability' or 'consistency' of the scores yielded by the scale. It seems reasonable to believe that two 'poetry' or two 'scientific' scales would yield scores that correlated to somewhat the same extent. Similar types of reading matter would probably correlate highly". And further on, "It appears then that ability in silent reading depends very largely on the nature of the passage read: a good reader with one type of material may not necessarily be a good reader with other material".

The above cited material is in striking agreement with some correlations which were worked out for the scores of sixteen pupils on the five types of materials used in the experiments reported in the present chapter. These correlations were worked out by a Spearman Rank-Difference method and then transmuted into Pearson's *r* by table (27, p. 492). Table 23 gives these correlations.

TABLE 23

CORRELATIONS BETWEEN PUPILS' RANKS IN FIVE DIFFERENT TESTS FOLLOWING A SINGLE READING

| Article 1 and | Article 2 | Test 1 | Test 2 | r | P.E. |
|---|---|---|---|---|---|
| Gov't of Switzerland | Peanuts | A | A | .61 | .0965 |
| Medieval Castles | Tuberculosis | B | B | .65 | .0871 |
| Rainbow | Peanuts | A | A | .73 | .0769 |
| Tuberculosis | Peanuts | B | A | .54 | .1131 |
| Rainbow | Tuberculosis | A | B | .54 | .1131 |
| Gov't of Switzerland | Tuberculosis | A | B | .50 | .1131 |
| Medieval Castles | Rainbow | B | A | .48 | .1261 |
| Gov't of Switzerland | Med. Castles | A | B | .48 | .1261 |
| Gov't of Switzerland | Rainbow | A | A | .45 | .1261 |
| Medieval Castles | Peanuts | B | A | .39 | .1373 |

The above table is read as follows:

"A" signifies that the controlled completion test was used in the measure of the effect of a single reading. "B" signifies that the question-statement test was used to measure the effect of a single reading.

It is to be noted that in the above tables there is a high correlation between the ability of the pupils to read the material "Government of Switzerland" and the material "Peanuts"; again, that "Peanuts" and "Rainbow" are highly correlated;

"Medieval Castles" and "Tuberculosis" are highly correlated; but "Government of Switzerland", which correlates with "Peanuts" does not correlate so highly with "Rainbow". All three of these selections are measured with the same type of test; hence it is probable that differences in ability to read the three types of material are shown. Again, it is seen that "Tuberculosis" and "Medieval Castles" correlate highly with each other but not so highly with the first three. Now these latter two selections are measured with a similar type of test. Of the two latter, "Tuberculosis" correlates more highly with the first three than "Medieval Castles". This is taken to mean that of the selections, the easiest to read is "Rainbow", the next "Peanuts", the next "Government of Switzerland", the next "Tuberculosis" and the last, "Medieval Castles". The gross scores on these articles also suggest this possibility. They are:

| *Material* | *Per cent reproduced* | *Character of Material* |
|---|---|---|
| Rainbow | 78 | Narrative |
| Peanuts | 47.5 | Agricultural |
| Gov't of Switz. | 43.9 | Government |
| Tuberculosis | 31.6 | Scientific |
| Med. Castles | 37.4 | Historical |

The lower score on "Tuberculosis" is attributed to the length of the material as contrasted with "Medieval Castles."

Now, since the first three of these scores are from tests of a similar character and the last two from tests of a different character, it is probable that the differences as brought out in the table of the scores and the failure of the tests to correlate highly are due to the nature of the tests as well as to the differences in difficulty of the materials, and that were all tests of the same type closer correlations would be shown.

These correlations are similar in their significance to those of Pressey and Pressey because they do bring out the fact that reading one type of material is not the same as reading another; it may be that reading ability, as suggested by these writers, is in need of more careful definition.

### *g. Further Conclusions from Experiments with Five Types of Material in a Sixth Grade*

1) The controlled completion is an easier type of test than the question-answer test. This is shown by the great differences in the amount reproduced by the different types of tests.
2) The effect of previous knowledge on materials of this type is consid-

erable; it is not safe to assume that pupils know nothing about an assignment which they are to read.
3) The relatively small scores resulting from a single reading emphasize the need of careful methods of assignment, or more practice in a single reading, more efficient methods of class instruction, if the child is to be efficient in his work.
4) The differences in the gross scores in experiments in which similar measuring devices are used suggest the need for a scientific determination of the relative difficulty of different materials on a large scale as a basis for devising efficient methods of dealing with the different subjects in the curriculum.

## 7. General Conclusions from the Present Chapter

The experimenter acknowledges that the conclusions which may be drawn from the experiments in this chapter are limited by the small number of cases in the latter experiments. However tentative the data may be the conclusions are very interesting, nevertheless, because of the questions which they open up and because of the general agreement which they show in regard to the effects of a single reading. Briefly, the general conclusions are as follows:

1) These experiments emphasize the fact that reading one type of material is probably not the same as reading another type theoretically similar in difficulty.
2 They suggest the desirability of further experimentation to determine the relative difficulty of different kinds of reading material.
3) They manifest the usual characteristics of silent reading experiments and are therefore probably as authentic in measuring the effect of reading as other silent reading tests which have been devised.
4) They bring out the loss which occurs as time elapses.
5) They open up the question of under-learning and its effect upon the reading of meaningful material under ordinary classroom conditions.
6) The rather small average scores resulting from a single reading indicate that this method of learning if not supplemented by more efficient devices may result in relatively small retention which decreases rapidly with the lapse of time, in some cases having apparently no effect after a month's time.
7) There are, however, some of the brighter children in all the tests, for whom a single reading is an efficient device, resulting in scores that closely approximate the total possible score, and remaining efficient after the lapse of time.

S.C.E.R.T., West Bengal
Library
Calcutta
25/3, B. C. Road

# CHAPTER VI

# EXPERIMENTS IN WHICH VARIABLES ARE CONTROLLED

In this chapter an account will be given of a series of experiments in which an effort was made to control and to estimate carefully the effect of several factors which had been uncontrolled in previous experiments. The series includes experiments given in seven different school systems involving about 500 pupils.

## 1. Purpose

When a child is given a selection to be read orally or silently, it is obvious:

1) That he already knows something about the selection, or he could not read it; that is to say, the elements of which the selection is made up must be so related to his experience that they have meaning for him. This previous knowledge is ordinarily inseparable from and a part of the reading complex, but it might be measured by a controlled experiment.

2) That previous experience may have brought him into contact with some of the facts in the selection which has been assigned for a new lesson.

3) That if one gives the child a test to determine what he already knows about the material, this test will have an effect upon his subsequent reading of the material.
   (a) In reviewing the facts known.
   (b) In motivating the reading, or showing him what to look for in the subsequent reading.

4) That if the previous test is the same as the test given for immediate recall, the effect of practice with the test itself may result in a higher score than would have been made had no test been given.

5) That if a test is given for the purpose of immediate recall, the effect of this test, as well as the effect of a single reading will be included in the score as the result of delayed recall.

It is clear that such uncontrolled factors as the above were present in the experiments previously reported and that, if the effect of a single reading, uninfluenced by other factors, is to be adequately measured or estimated as accurately as possible they must be isolated and measured. In the following pages an account is given of the manner in which the experimenter attempted to measure or estimate these various factors.

## 2. MATERIALS

### *a. Reading Materials*

Two types of reading materials were used in this series of experiments: the first, an article adapted from "The Living Races of Mankind" by H. N. Hutchinson and rewritten by the experimenter; and the second, the article, "On the Use of Abbreviations" by Jean Sherwood Rankin.

The first of the above articles is called in this chapter, "The Admiralty Islanders." It is an article of 1000 words, of interesting, very detailed material designed to approximate in difficulty such descriptions of different peoples as are commonly found in intermediate and advanced geographies, or in simple books of travel, but to contain facts about a people with whom the pupils are unlikely to be familiar. The volume from which the article was taken is a large book in a remote corner of the University library, so unwieldy as to be little used. It is therefore quite improbable that pupils in the elementary school would ever come into contact with the particular facts included in this selection. For this reason, the material is particularly well adapted to the experiment. It represents material such as the pupil might be required to read when coming into contact for the first time with a new chapter in his geography.

The second article is a highly factual selection from a language book. In this chapter it is called, "On the Use of Abbrevitions." The article is 1200 words in length, factual, practical, and interesting, but very difficult. It is not so well adapted to a single silent reading as the first article described above, but it is material with which pupils come into contact in the school; therefore, the value of a single reading on this type of reading should be measured.

### *b. Test Material*

The tests used in the experiments described in this chapter are of two types: the first, a controlled completion test of 132 points used to test the effect of a single reading of the material, "Admiralty Islander"; and the second, a *Yes and No* test of 100 points, containing fifty pair answers.

The controlled completion test used in this series is very exhaustive. It involves a point for practically every idea in the whole selection. It makes no effort to divide main ideas from subordinate ideas, but designs to find out whether or not the pupil is able to recognize the fact that he has just read about each point. It is probable that the test is so long that for practical schoolroom testing it should be modified. The labor of checking this test is also so great that it is difficult to measure large numbers of children with it.

The *Yes and No* test used in the second of the experiments is designed to eliminate chance or guessing the answer. The pupil is required to answer both *Yes* and *No* to each and every point in the test; that is, every point is stated in both a negative and a positive form. If a child answers the point correctly in both its negative and its positive form, it is quite certain that he understands it. This test is very long and laborious to score and could not be used where economy of time in measuring immediate results is a factor, but it is well adapted to a refined experiment. It was made by reworking a test used in a set of preliminary experiments. The experimenter analyzed a set of papers in which this short test was employed and determined which exercises were answered by all or nearly all the pupils. He eliminated these answers and made a new test of the remaining points in the article. This new test consisted of 100 points, fifty of which were positive statements of facts and fifty of which were negative statements of the same facts. The child is required to answer each point both negatively and positively. Chance answers are thus very greatly decreased. The paired answers are arranged in such a haphazard way that there is little likelihood of the child comprehending the scheme and answering the questions by chance. Since, in addition to these precautions to rule out chance, the children were instructed not to answer the question unless they felt that they knew the answer, it is probable that the test eliminated this factor.

## 3. SCORING

The tests were both scored in the same manner as those used in the preceding experiments. Each correct answer was scored one point and each incorrect answer given nothing. The method of arriving at the final score was to add the total number of points correct and divide this sum by the total possible points in the selection. The answer is thus expressed as a per cent of the total possible score. A discussion of the validity of this method of scoring will be found in a succeeding chapter.

### 4. Method of Conducting Experiments

The method of conducting the experiments in this series is in most respects the same as in preceding experiments. Each teacher was given a carefully worked-out set of directions and a prepared outline of the experiment. The experiments on the material "Admiralty Islanders" were conducted by the writer and two other trained experimenters; the experiment with the materials "On Abbreviations" was under the direction of Superintendent Ray Latham of Ida Grove, Iowa, a trained student who is competent in every way to conduct an experiment of this sort. He was provided with the careful instructions reproduced herewith.

### INSTRUCTIONS TO TEACHERS
### VALUE OF A SINGLE READING

by

G. A. YOAKAM

The object of this experiment is to measure the effect of a single reading on content material such as occurs in geography and history. The experiment is scientifically set up. Your coöperation is asked in seeing that the following conditions are kept uniform in so far as is possible.

1. Pupils should not discuss the material during the interval between tests. Two tests will be given to groups A and B. Group C will be given one test only.
2. That no class discussion of the material take place until the experiment is finished.
3. That pupils in groups A, B and C, do not discuss the material, since such discussion would go far to invalidate the conclusions which it is hoped will be drawn.

The general plan of the experiments is as follows:

| Group A | Group B | Group C |
|---|---|---|
| Test | — | — |
| Read | Read | Read |
| Test | Test | — |
| | Interval of two or three weeks | |
| Test | Test | Test |

The need for a measurement of the effect upon this type of material of a single reading is very great. Until we know the effect of a single reading, we cannot have any basis for argument as to special methods of study. This experiment fills a gap in our knowledge of the effect of study which has long been needed. Its general position is as follows: It bridges the gap between no knowledge at all and that knowledge which the child possesses in the ordinary school, when he comes to the recitation. Until we know how great is the effect of the reading in the child's study period—commonly a single reading—we have no basis for the measurement of what has happened during the recitation period.

Your coöperation in making this experiment effective is greatly desired.

The chief difference between the method of conducting this series of experiments and that of previous experiments is in the division of each grade into three groups for the purpose of using the results to establish corrections for the effect of previous knowledge, the effect of the motivation of the initial test upon the subsequent reading of the article, the effect of the practice with the test upon immediate recall, and the effect of the practice with the test upon delayed recall. The plan of procedure is best shown by the outline given in the instructions for teachers on the preceding page. It is seen that each grade was divided into three groups, Group A, Group B, and Group C. The procedure in connection with each group differed from that with every other. Thus Group A was tested, read the article and was tested again; Group B read the article once and was tested; Group C read the article but was not tested until the time for a delayed recall by the other groups. If the accurate division into groups is made, it is clear that the scores made by Group A in the initial test will represent previous knowledge of the material, plus whatever chance there may be to guess the answers of the test; that the scores made on the immediate recall test will not be scores of a single reading alone but of a single reading plus practice with the test, plus the effect of motivation of the previous test upon the single reading. It is also quite as obvious that the immediate recall scores of Group B which has not had an initial test will not be affected by practice with the test, or by the effect of the motivation caused by the use of the test. It may thus be possible to correct the scores of Group B for the effect of previous knowledge of the material by subtracting from the immediate recall scores of Group B the initial test scores of Group A. Further, it is clear that Group C, which has had no immediate recall, will not be influenced by the effect of the repetition of the test after a single reading and that the scores made by this group will represent the effect of a single reading uninfluenced by the immediate recall which has been given to Group B. It is therefore possible to use the scores made by Group C to correct the delayed recall scores of Group B. for the effect which the practice with the test incident to immediate recall has had upon delayed recall. Thus, by the use of the corrections arrived at in this manner, it may be possible to make a fairly reliable estimate of that effect of a single reading uninfluenced by the other factors than reading itself, which were present in the previous experiments.

The success of the above procedure is, however, contingent upon the manner in which the groups are divided and the treatment made of the scores of the various groups. In searching for a method of dividing the grades into groups, the writer thought of using the intelligence quotients of the various members as a basis of division, but this appeared undesirable since intelligence and reading ability, while related, are not necessarily found together. The common method of dividing grades into groups for experimental purposes has been by means of reading tests or by the judgment of teachers. Hoover (28) used the latter method in his significant series of experiments reported in the *Twentieth Year Book*. Later he checked his division by means of scores made on Monroe's Silent Reading Test. He says, "After these tests were given, the attempt was made to divide the pupils into two groups of equal size and mental attainments. In making this division the advice and assistance of the superintendent of schools was sought. This method of division seemed advisable, as it was desirable to begin without delay rather than to wait until all the papers had been scored and to make the division then upon the basis of the results thus found. The superintendent's judgment in this matter was exceedingly accurate, as will be noted by referring to data in Table 1. A better division of the children could not have been realized even if the division had been made upon the results of tests." This statement is cited because in the following experiments there was a division of grades into groups (1) by means of the Monroe Silent Reading Tests, and (2) by means of the judgment of teachers. The teacher's judgment has been questioned as a means of dividing grades into groups. The writer is of the opinion that it is possible by this means to get two fairly accurate groups of approximately equal ability and the results of the above experiment and the data following seem to prove the correctness of the opinion. In part of the experiments reading tests were unavailable for division of groups. Accordingly the teachers ranked the pupils into groups of good, medium, and poor ability. The experimenter then put by chance selection roughly one-fourth of the pupils into Group A, one-half into Group B, and one-fourth into Group C. In another experiment with pupils in three grades in two schools, pupils were ranked by means of the Monroe Silent Reading Tests and then divided into three equal groups.

Any division of grades into groups is more or less inaccurate because of the variation in the judgment of teachers and because of the shifting of pupils on successive reading tests. This was brought out in the following experiments: Pupils who are ranked as of highest ability in the reading test, often did not perform as well as those of less ability when tested by the experimental material. The performance of the pupils with the material seems, however, to have justified the division and the treatment of the grade curves resulting from these data possibly iron out some of the irregularities in division due to the foregoing factors. Moreover, with a sufficient number of pupils, these irregularities are probably compensated for. In the following tables are represented the scores made by pupils in five schools in two different states, The following shows the total number.

TOTAL NUMBER OF CASES TESTED

| | VIII | VII | VI | V | IV |
|---|---|---|---|---|---|
| Group A | 35 | 25 | 48 | 40 | 8 |
| Group B | 47 | 37 | 72 | 32 | 17 |
| Group C | 23 | 23 | 43 | 27 | 8 |
| Totals | 95 | 85 | 63 | 99 | 33 |

The total number tested is 47.5, of which 291 are used in making the corections for previous knowledge, motivation, and practice with the test. The scores for pupils in grades eight, seven, and five only are used in making corrections because in these grades the division into groups seems to have been most successful. The inequality of groups is due to the fact that pupils in some groups were more numerous than in others, as stated above, owing to the manner of division and also to the fact that some pupils were not present at the time of the delayed test and thus were not included in the results.

The performances of the pupils with the material seems, however, to have justified the selection and to have resulted in a somewhat reliable correction for the various factors which the foregoing procedure was designed to measure. Since, in most of the grades, the corrections are based on the scores from more than one school, it is thought that the effect of the selection of pupils in one school where some irregularity might have occurred, is probably balanced to some extent by the selection in another.

The resulting division is probably fairly representative; therefore, the resulting averages which are used for corrections, may be used with some confidence in estimating the effect of previous knowledge, practice with the test, and the effect of motivation in reading.

## 5. Results of Experiment with the Article "The Admiralty Islanders"

### *a. Introduction*

The experiment with geographical material, "The Admiralty Islanders," involved approximately 373 pupils of six schools in grades four to eight inclusive. Five of these schools were ordinary graded schools and the third was the University Elementary School. Pupils in the latter school were so well trained in silent reading that their scores were invariably higher on the average than those of the same grade in other schools. This probably tended to raise the averages in the grades for which their scores were used.

The following sections summarize the raw data in the form of frequency tables.

### *b. Interpretation of Data Without Regard to Grades*

The scores are treated at first without regard to grade. This is perhaps wise since it is sufficient for the purposes of these experiments to estimate the amount of correction for previous knowledge, motivation, and practice with the test in general terms for the upper grades of the elementary school. Averages are always liable to cover up truths and to leave out of account the differences between individuals, the range of abilities found and the amount of deviation from the mean. Table 24 gives the initial, immediate and delayed recall scores for Group A, the immediate and delayed recall scores for Group B, and the delayed recall scores for Group C. The scores are consistent for the groups. In tabulating the original scores it was observed that in range and central tendency the scores for Group A were consistently larger than those for Group B, on both immediate and delayed recall, and that the scores for Group C on delayed recall were consistently smaller in range and central tendency than those for Groups A and B. This is attributed to the influence of the reading since the groups could not have been consistently divided under this plan so that Group A was superior to Group B, and Group B superio to Group C. In fact, as the results show, this was not the case.

TABLE 24

DISTRIBUTION OF THE SCORES OF ALL GROUPS WITHOUT REGARD TO GRADE

Grades 5, 7, and 8
"Admiralty Islanders"
291 Pupils

| Class Interval | | | |
|---|---|---|---|
| | GROUP A | | |
| | *Initial* | *Immediate* | *Delayed* |
| 80–89.9 | | 1 | |
| 70–79.9 | | 11 | |
| 60–69.9 | | 12 | 4 |
| 50–59.9 | | 22 | 10 |
| 40–49.9 | | 13 | 14 |
| 30–39.9 | 1 | 8 | 17 |
| 20–29.9 | 24 | 13 | 11 |
| 10–19.9 | 45 | 11 | 11 |
| 0– 9.9 | 25 | 4 | 2 |
| Total | 100 | 95 | 68 |
| Median | 15.6 | 48.8 | 38.8 |
| Q | 5 | 17 | 12.1 |
| | GROUP B | | |
| 80–89.9 | | | |
| 70–79.9 | | 1 | |
| 60–69.9 | | 11 | 1 |
| 50–59.9 | | 14 | 3 |
| 40–49.9 | | 28 | 18 |
| 30–39.9 | | 27 | 23 |
| 20–29.9 | | 19 | 19 |
| 10–19.9 | | 11 | 15 |
| 0–19.9 | | 7 | 9 |
| Totals | | 118 | 88 |
| Medians | | 39 | 30.4 |
| Q | | 11.2 | 10.7 |
| | GROUP C | | |
| 40–49.9 | | | 1 |
| 30–39.9 | | | 4 |
| 20–29.9 | | | 32 |
| 10–19.9 | | | 28 |
| 0– 9.9 | | | 8 |
| Total | | | 73 |
| Median | | | 20.1 |
| Q | | | 6.3 |

Table 24 should be read as follows; in grades five, seven, and eight one pupil made a score of 30–39.9 on the initial test, or, that is to say, he was able to do correctly approximately 35% of the possible exercises in the test before a single reading, etc. The remainder of the scores are read in a similar manner.

The following interesting data appear in the table:

1) Previous knowledge or ability to do the test without reading the

article ranges from 0 to 39 per cent of the total ideas in the article. The median for these grades is 15.6 per cent of the total ideas.

2) Ability to do the test after a single reading preceded by a test ranges from 0 per cent to 89.9 per cent of the total ideas in the article. The median for these grades is 48.8 per cent of the total ideas or slightly less than half.

3) Delayed recall in Group A after an initial test and an immediate recall test, 20 days elapsing, ranges from 10 per cent to 69.9 per cent. The median is 38.8 per cent. There has been a loss in 20 days time of 10 per cent of the total ideas in the article.

4) The scores of Group B on immediate recall are in range from 0 per cent to 79.9 per cent; the median for this group is 39 per cent. This is 9.8 per cent, on the average, less than the scores of Group A which benefited from an initial test which probably motivated the reading. The effect of this motivation and practice with the test is probably somewhere near 10 per cent.

5) On delayed recall the scores of Group B range from 0 per cent to 69.9 per cent. The median of the averages is 30.4 per cent. The loss in this group due to the lapse of 20 days is 8.6 per cent or approximately the same as that for Group A.

6) Group C on delayed recall ranges in ability from 0 per cent of the total ideas to 49.9 per cent. The differences between the delayed recall scores of Groups A and B respectively and Group C are 18.7 per cent and 10.3 per cent of the total ideas in the article.

7) Q or the quartile deviation is small before reading, is greatly increased after reading, and grows smaller again after 20 days have elapsed; that is to say, the range of ability to do the test increases on immediate recall and decreases slightly on delayed recall.

### c. *Establishing Corrections for Variables*

From the above data it appears that the variable factors (previous knowledge, motivation, and effect of practice with the test, together with a single reading) are probably operating in the following manner. The corrections are computed by using the medians of the averages for Groups A, B, and C as explained on page 56.

| | |
|---|---|
| Previous knowledge and ability to do the test due to guessing or natural logical thinking (Group A—Initial) | 15.6% |
| Effect of motivation of the initial test (A immediate—B immediate) | 9.8% |
| Effect of single reading (Immediate score B—Initial score A) | 24.5% |
| Effect of single reading (Immediate A—Initial A—Effect of motivation) | 25.4% |
| Effect of immediate recall test on delayed recall (Delayed B—Delayed C) | 10.3% |
| Effect of single reading on delayed recall (B—Initial A—Delayed C) | 4.5% |

These corrections, to be sure, are at the best speculative but they at least make it appear probable that the effect of a single reading is on the average in these grades to give the reader an ability to do 25 per cent more on immediate recall, and only about 4.5 per cent more on delayed recall after 20 days.

### *d. Conclusions*

The scores given in the preceding table and the corrections computed on those scores lead to the following tentative conclusions in regard to the effect of a single reading as a device for efficient work in school. Although indispensable, it is shown to be relatively inefficient.

1) A single reading uninfluenced by previous knowledge in grades 5, 7, and 8 results in, on the average, only 25 per cent of the total ideas in an article on immediate recall.

2) On delayed recall after 20 days a single reading alone would probably give very little ability to do this test beyond that already possessed by the pupils because of previous knowledge of the material.

3) An initial test given for the purpose of testing out previous knowledge is an efficient device for study because it motivates and directs reading.

4) For purposes of delayed recall repeated testing is more efficient than the single, though indispensable, contact with the material itself. This is in accordance with the laws of learning. Frequent repetition of an idea secures permanent retention.

5) The ability to read and reproduce after a single reading varies greatly among individuals. The range is from nothing to a maximum of 89.9 per cent in these grades, but the high scores are relatively few. The upper quartile on immediate recall in Group B is 49.2 per cent and the lower 26 per cent, while the median is 39 per cent. That is to say, if children in these grades read this material a single time without having it motivated by a problem or some other purpose only one-fourth of them will get as much as 50 per cent of the total ideas in the article. This argues strongly for the proper emphasis upon directed study. This argument will be supported later by data which show the overlapping among these grades in this ability and emphasize the large individual differences of children in this skill.

### *e. A study of these data when distributed by grades.*

In order to emphasize grade differences and to give an idea of overlapping of reading ability in respect to this material as shown in the schools and grades investigated, the same data as reported in the previous section are now distributed by grades. Some data gathered in grades 6 and 4 are also introduced although these data are not believed to be so reliable on account of the unequal numbers in the groups. The preliminary report of this material was contributed by the writer to the *Twentieth Year Book of the National Society for the Study of Education* and may be found there in some detail (29).

Distribution tables for the various groups and grades follow.

TABLE 25

DISTRIBUTION OF SCORES OF GROUP A

Initial Test for Previous Knowledge by Grades and Immediate Recall
"The Admiralty Islanders"

| Class Interval | Grades<br>VIII | VII | VI | V | IV |
|---|---|---|---|---|---|
| 30–39.9 | 1 | | | | |
| 20–29.9 | 15 | 9 | 3 | 5 | |
| 10–19.9 | 16 | 14 | 21 | 15 | 2 |
| 0– 9.9 | 3 | 2 | 24 | 20 | 6 |
| Totals | 35 | 25 | 48 | 40 | 8 |
| Medians | 19.3 | 18.2 | 12.3 | 11.2 | 6 |
| Q | 5.2 | 5.1 | 5.4 | 5.8 | 3.S |
| | | *Immediate Recall* | | | |
| 80–89.9 | 1 | | | | |
| 70–79.9 | 8 | 2 | 3 | 1 | |
| 60–69.9 | 6 | 3 | 2 | 3 | |
| 50–59.9 | 10 | 8 | 11 | 4 | |
| 40–49.9 | 6 | 4 | 9 | 3 | |
| 30–39.9 | 11 | 2 | 12 | 5 | 2 |
| 20–29.9 | 1 | 2 | 2 | 10 | 2 |
| 10–19.9 | 1 | 1 | 4 | 9 | 2 |
| 0– 9.9 | 1 | | | 4 | 2 |
| Totals | 35 | 25 | 48 | 40 | 8 |
| Medians | 59.5 | 52.5 | 47.5 | 36.5 | 19. |
| Q | 11.3 | 7.2 | 14.3 | 13.4 | 10. |

The following table gives scores for Group A, delayed recall.

TABLE 26

DISTRIBUTION OF SCORES OF GROUP A BY GRADES

DELAYED RECALL

20 days

"Admiralty Islanders"

| Class Interval | Grades<br>VIII | VII | VI | V | IV |
|---|---|---|---|---|---|
| 60–69.9 | 4 | | | | |
| 50–59.9 | 6 | 4 | 1 | | |
| 40–49.9 | 4 | 9 | 1 | 1 | |
| 30–39.9 | 6 | 6 | 2 | 5 | 1 |
| 20–29.9 | 2 | 2 | 2 | 7 | 2 |
| 10–19.9 | | 2 | 2 | 4 | 3 |
| 0– 9.9 | | | | 2 | 2 |
| Totals | 22 | 23 | 8 | 21 | 8 |
| Medians | 47.5 | 41.6 | 30. | 26.4 | 16.6 |
| Q | 10.8 | 7.8 | 10. | 6.2 | 7.5 |

The above tables are read as follows: on delayed recall 4 pupils in grade 8 made scores of from 60-69.9 pe rcent of the possible score, etc. The median for grade of the averages for this grade is 47.5 per cent of the total possible score etc.

The reader's attention is called to the following interpretation of the above table:

1) The scores on the initial test in all grades are consistently small; there is progress made from grade to grade and there is a large amount of overlapping in these grades. On this test the amount of overlapping may be expressed in terms of per cent as follows:

36% of the pupils in Grade 7 overlap the upper 50% of Grade 8.

6% of the pupils in Grade 6 overlap the upper 50% of Grade 8.

12.5% of the pupils in Grade 5 overlap the upper 50% of Grade 8.

2) Scores on immediate recall are greatly increased consistently throughout the grades. The only uncontrolled variable is the single reading preceded by the test. This measures the amount of a single reading.

3) Overlapping in these grades on immediate recall is just as significant.

20% of the pupils in Grade 7 overlap the upper 50% of Grade 8

10% of the pupils in Grade 6 overlap the upper 50% of Grade 8

10% of the pupils in Grade 5 overlap the upper 50% of Grade 8

This means that grading of reading ability in these grades is very rough indeed and that pupils of at least the four upper grades do not differ so greatly that what may be said of one grade will not be roughly true of the others.

4) On delayed recall the scores in all grades consistently decrease, due to the lapse of 20 days. The loss is greater in the lower than in the upper grades, suggesting that while pupils in the middle grades have the skill pretty well perfected, they do not remember as long. The time taken to do the test in their case is greater also.

5) The amount of overlapping on delayed recall is consider-

able but not so great as on immediate recall. It may be expressed as follows:

91% of the pupils in Grade 7 overlap the lowest pupils in Grade 8
75% of the pupils in Grade 6 overlap the lowest pupils in Grade 8
80% of the pupils in Grade 5 overlap the lowest pupils in Grade 8
75% of the pupils in Grade 4 overlap the lowest pupils in Grade 8

The scores of Group B on immediate and delayed recall will be shown next. It will be remembered that Group B had no initial test.

TABLE 27

SCORES OF GROUP B DISTRIBUTED BY GRADES
IMMEDIATE AND DELAYED RECALL
"Admiralty Islanders"

Immediate Recall

| | | | Grades | | |
|---|---|---|---|---|---|
| Class Interval | VIII | VII | VI | V | IV |
| 70–79.9 | 1 | | | | |
| 60–69.9 | 7 | 2 | 1 | | |
| 50–59.9 | 8 | 5 | 5 | 1 | |
| 40–49.9 | 16 | 8 | 12 | 4 | 1 |
| 30–39.9 | 8 | 11 | 17 | 8 | |
| 20–29.9 | 6 | 7 | 20 | 6 | 3 |
| 10–19.9 | 1 | 3 | 7 | 7 | 10 |
| 0– 9.9 | | 1 | 10 | 6 | 3 |
| Totals | 47 | 37 | 72 | 32 | 17 |
| Medians | 45.3 | 37.7 | 29.5 | 25. | 15.5 |
| Q | 13.9 | 9.9 | 9.7 | 12.2 | 9. |
| | | Delayed Recall (20 Days) | | | |
| 60–69.9 | 1 | | | | |
| 50–59.9 | 2 | 1 | | | |
| 40–49.9 | 9 | 7 | 4 | 1 | |
| 30–39.9 | 7 | 9 | 4 | 7 | 1 |
| 20–29.9 | 6 | 8 | 10 | 5 | 1 |
| 10–19.9 | 1 | 5 | 2 | 10 | 11 |
| 0– 9.9 | | | | | 4 |
| Totals | 26 | 30 | 20 | 32 | 17 |
| Medians | 38.6 | 32.2 | 28. | 27. | 13.6 |
| Q | 8.1 | 8.8 | 7.3 | 10.6 | 3.8 |

The reason for the difference in the size of the groups in the above immediate and delayed recall scores is that scores are contained in the immediate recall from one school in which, through a slip, delayed recall scores were not given. If the selection of pupils in these groups is made so that pupils of somewhat equal ability are chosen, the effect of adding more scores to the group is to make the median somewhat more reliable. For instance, since the pupils in grade six, immediate recall are representative of six schools they are regarded as reliable for grade six on this phase of the test. The delayed recall scores, however, are not so reliable.

Before pointing out the results shown in the above tables the delayed recall scores for Group C which had no immediate recall nor any initial test but a single reading only are given in Table 28.

TABLE 28

SCORES OF GROUP C DISTRIBUTED BY GRADES

DELAYED RECALL

"Admiralty Islanders"

| Class Interval | Grades<br>VIII | VII | VI | V | IV |
|---|---|---|---|---|---|
| 40–49.9 | | 1 | 1 | | |
| 30–39.9 | 4 | | 3 | | |
| 20–29.9 | 12 | 12 | 15 | 8 | 1 |
| 10–19.9 | 7 | 8 | 16 | 13 | 3 |
| 0– 9.9 | | 2 | 8 | 6 | 4 |
| Totals | 23 | 23 | 43 | 27 | 8 |
| Medians | 23.7 | 21.2 | 18.4 | 15.7 | 10. |
| Q | 5.2 | 5.7 | 7.4 | 7.3 | 5.8 |

The foregoing tables are read in the same manner as those immediately preceding. From a study of the data given in Tables 27 and 28, the following results seem to be shown:

1) On immediate recall the scores of Group B in the different grades are consistently smaller than the scores for Group A on the same test. They decrease from grade to grade as shown in the scores.

2) The amount due to a single reading ranges from 45.5 per cent in Grade 8 to 15.5 per cent in Grade 4. Part of this is no doubt due to previous knowledge so that the amount due to reading itself is comparatively small. This amount will be computed when corrections are made in the following pages.

3) The scores of Group B on delayed recall show a marked ability of the pupils to perform the test after a lapse of 20 days, but there is the usual overlapping and wide range in individual abilities. The averages are consistently lower than on immediate recall. Progress in the ability to do this test after a single reading and an immediate recall test is steady from grade to

grade. When corrected for previous knowledge, however, the amounts attributed to a single reading become surprisingly small.

4) The scores of Group C on delayed recall show that the ability to do this test after a single reading and a lapse of 20 days is very slight. It does not exceed to any marked degree the ability of Group A to do the test before any reading occurred. It is therefore probable that the single reading without immediate recall has left very little impression upon the minds of these pupils although some slight effect is discernible in the delayed recall scores.

5) In all these tables the dispersion of the groups around the central tendency is shown to be greater after a single reading than it is 20 days later. This suggests that the single reading changed the amount of knowledge of this material as to cause those who were good readers greatly to enlarge their knowledge of the material according to their abilities but that when time intervened to cause forgetting, the individual differences in the knowledge of the groups naturally decreased. Pupils who read well had much to forget; those who read poorly less. Hence the the interval caused the knowledge of the groups to become more nearly as it was before any reading had taken place.

### *f. Conclusions from grade distributions*

From a study of the data given in the grade distributions the following inferences are drawn.

1) Ability to reproduce this material after a single reading increases in a somewhat regular manner from grade to grade.

2) There is marked overlapping in this ability as shown in these grades so that the generalizations made in Section 5*d* of this chapter seem to be roughly true for pupils in these grades as a whole.

3) Ability to reproduce material after a single reading is relatively slight when variable factors are taken into consideration.

4) The division of grades into groups and the attempt to measure the factors of previous knowledge, motivation, and practice with the test seem to have been successful. Differences in grades and groups are clearly and consistently shown.

5) The data given in the foregoing tables may be used to esti-

mate the effect of the variables. This will be done in the following pages.

### *g. Correcting for the Variables*

In general the tables represent very clearly the results of the measurement of relatively increased power to perform the tests after a single reading as measured by immediate recall and a decrease in this same power after a lapse of 20 days. It is probable that the medians would shift if a larger number of cases were taken; but from experience with the tests in several school systems the experimenter is of the opinion that this shift would not be great. It might be calculated by statistical treatment of the data but this is thought neither necessary nor desirable. The data show only *what they show* for these schools and grades and under these conditions. Changed conditions with more widely selected pupils would alter the amounts of these medians but not the general significance of the facts, nor the relative amounts due to previous knowledge, the effect of a single reading, and the effect of motivation and practice with the tests. The general significance of the experiment would not be changed.

The method of making the correction for the variables (previous knowledge, effect of motivation of the initial test in Group A, the single reading, and the effect of the repetition of the tests upon delayed recall) is, as was explained earlier in this chapter, to use the median of the average scores made by the three groups in each grade, Groups A, B, and C interchangeably, the scores of one group being subtracted or added to the scores of another to account for the effect of the variables. These corrections are given only for speculative purposes here. The writer does not claim for them any fine degree of accuracy.

The first correction made is to arrive at the effect of a single reading plus motivation due to the initial test in the scores of Group A by subtracting from the immediate recall scores the scores made in the initial test.

TABLE 29

CORRECTION OF SCORES OF GROUP A FOR THE EFFECT OF PREVIOUS KNOWLEDGE

"The Admiralty Islanders"

| Grades | VIII | VII | VI | V | IV |
|---|---|---|---|---|---|
| Immediate recall—Group A | 59.5 | 52.5 | 47.5 | 36.5 | 19 |
| Initial scores—Group A | 19.3 | 18.2 | 12.3 | 11.2 | 6 |
| Effect of single reading plus motivation | 40.2 | 34.2 | 35.2 | 25.3 | 13 |

It is obvious that when the initial scores are subtracted from the immediate recall scores that there is still a factor affecting the result—the effect of the motivation caused by the initial test. The only way to arrive at an approximate estimate of this effect is to use the scores of Group B, corrected for previous knowledge, to correct the scores of Group A for the effect of this motivation. Table 30 gives the scores of Group B corrected for the amount of previous knowledge. One way to arrive at this correction is to assume that the previous knowledge of Group A and Group B is approximately the same and to subtract the initial scores of Group A from the immediate recall scores of Group B.

TABLE 30

CORRECTION OF THE SCORES OF GROUP B FOR THE EFFECT OF PREVIOUS KNOWLEDGE

"The Admiralty Islanders"

| Grades | VIII | VII | VI | V | IV |
|---|---|---|---|---|---|
| Immediate recall—Gr. B | 45.3 | 37.7 | 29.5 | 25. | 15.5 |
| Initial score—Gr. A | 19.3 | 18.2 | 12.3 | 11.2 | 6. |
| Effect of single reading | 26.0 | 19.5 | 17.2 | 13.8 | 9.5 |

It is seen from the above table that when the correction for the effect of previous knowledge is made the resulting scores, due to the single reading, are comparatively small, and the largest difference between grades is between Grades 4 and 5, and between Grades 7 and 8.

To correct the score of Group A for the effect of motivation of the test upon the reading, we subtract the single reading scores of Group B from the corrected scores of Group A as shown in Table 31, and we have the effect of motivation, of the supperior ability of Group A, or of both.

TABLE 31

CORRECTION OF SCORES OF GROUP A FOR EFFECT OF MOTIVATION OF INITIAL TEST

"The Admiralty Islanders"

| Grades | VIII | VII | VI | V | IV |
|---|---|---|---|---|---|
| Group A corrected for previous knowledge | 40.2 | 34.3 | 35.2 | 25.3 | 13. |
| Group B corrected for previous knowledge | 26. | 19.5 | 17.2 | 13.8 | 9.5 |
| Effect of motivation | 14.2 | 14.8 | 18.0 | 11.5 | 3.5 |

It is recognized that the series of averages in the above table labelled, "Effect of Motivation," may be due to differences in the groups as well as to the effect of motivation, but it is unlikely that the differences in the groups account for it wholly. The fact that Group A consistently has a larger score in the series of tables would tend to show that this is due to the nature of the method used with this group rather than to the effect of a more selective group.

The next step is to correct the delayed recall scores of Group B for the effect of the practice in recall due to the immediate recall test. To do this the delayed recall scores of Group C are subtracted from the delayed recall scores of Group B, since Group C had no immediate recall; their scores should represent what Group B would have done had they not been given an immediate recall.

### TABLE 32

CORRECTION OF SCORES OF GROUP B DELAYED RECALL FOR EFFECT OF IMMEDIATE RECALL

"The Admiralty Islanders"

| Grades | VIII | VII | VI | V | IV |
|---|---|---|---|---|---|
| Group B delayed recall | 38.6 | 32.2 | 28. | 27. | 17. |
| Group C delayed recall | 23.7 | 21.2 | 18.4 | 15.7 | 10. |
| Effect of immediate recall | 14.9 | 11. | 9.6 | 11.3 | 7. |

By proceeding in the same manner and correcting for the effect of the various tests upon subsequent tests, the effect of reading upon tests, and the like, it is possible to arrive at a final table which gives the effect of the various factors in summary form. These figures are given in the following tables:

### TABLE 33

EFFECT OF VARIOUS FACTORS AS MEASURED BY INITIAL TEST AND IMMEDIATE RECALL

"The Admiralty Islanders"

| | Grades | | | | |
|---|---|---|---|---|---|
| | VIII | VII | VI | V | IV |
| Effect of previous knowledge | 19.3 | 18.2 | 12.3 | 11.2 | 6. |
| Effect of single reading | 26. | 19.5 | 17.2 | 13.8 | 9.5 |
| Effect of motivation | 14.2 | 14.8 | 18.5 | 11.5 | 3.5 |

### TABLE 34

EFFECT OF VARIOUS FACTORS AS MEASURED BY DELAYED RECALL

"The Admiralty Islanders"

| | Grades | | | | |
|---|---|---|---|---|---|
| | VIII | VII | VI | V | IV |
| Effect of single reading | 4.4 | 3.0 | 6.1 | 4.5 | 6.2 |
| Effect of immediate recall | 4.9 | 11.0 | 9.6 | 11.3 | 3.6 |
| Effect of reading and two tests | 23.9 | 20.4 | 11.6 | 10.7 | 6.6 |
| Effect of reading and one test | 14.9 | 11.0 | 9.6 | 11.3 | 3.6 |
| Effect of two tests | 34.2 | 29.2 | 21.9 | 22.5 | 9.4 |

In the above tables the scores should be read as per cents of the total possible score.

Conclusions drawn from these tables will be deferred to the end of this chapter.

The remaining data which are gathered from this series of experiments show the amount of loss due to the lapse of the 20-day interval between the immediate and the delayed recall tests. The differences between the loss by Group A and the loss by Group B are seen to be slight.

TABLE 35

LOSS DUE TO THE LAPSE OF TIME BETWEEN IMMEDIATE AND DELAYED RECALL TESTS

"The Admiralty Islanders"

| | Grades | | | | |
|---|---|---|---|---|---|
| Grades | VIII | VII | VI | V | IV |
| Group A Loss | 12. | 10.9 | 17.5 | 10.1 | 2.4 |
| Group B Loss | 7.7 | 5.5 | 1.5 | –2 | 1.9 |

It is seen by the table that the loss in each group after the lapse of 20 days is comparatively slight. Reference to the tables will show, however, that Group A maintains its relative superiority over Group B on delayed as well as on immediate recall, when measured by gross score. One might naturally expect that Group B having had one less contact with the test than Group A would have forgotten more rapidly. But perhaps Group B loses less because there is less to lose.

### 6. Experiment with the Material "Abbreviations"

Before summing up the results of this chapter it will be necessary to give a brief summary of the results of an experiment with the material "On the Use of Abbreviations" in Grade 8 of the Ida Grove Public Schools. This experiment was conducted in the same manner as the preceding. Pupils were divided by the teacher's judgment into three groups of approximately equal reading ability. The type of test used was a *Yes and No* Test of 100 paired questions. The following table gives the scores of Group A, Group B, and Group C. The total number of cases is 26.

The reader will note that the results of this experiment with different material and a different test are much the same as the previous experiments reported in this study. There is considerable previous knowledge and the effect of the motivation of the initial test is marked. The only difference to be noted in the above table is the fact that there is apparent reminiscence taking place in the delayed recall; for after 6 days time the score is larger than after the immediate recall. The amount of previous

TABLE 36

SCORES OF EIGHTH GRADE ON "ABBREVIATIONS"

Per Cent of Paired Answers Correct

| Group A | Initial | Immediate Recall | Delayed Recall |
|---|---|---|---|
| 1 | 42 | 70 | 66 |
| 2 | 34 | 54 | 62 |
| 3 | 32 | 50 | 54 |
| 4 | 18 | 50 | 54 |
| 5 | 26 | 38 | 34 |
| 6 | 16 | 34 | 32 |
| 7 | 30 | 28 | 28 |
| Averages | 28 | 46 | 47 |
| P. E. of Av. | 2.08 | 3.34 | 3.28 |
| Group B | | | |
| Pupil | | | |
| 1 | | 64 | |
| 2 | | 52 | 66 |
| 3 | | 52 | 58 |
| 4 | | 44 | 42 |
| 5 | | 36 | 48 |
| 6 | | 36 | |
| 7 | | 34 | 36 |
| 8 | | 32 | 50 |
| 9 | | 28 | 32 |
| 10 | | 28 | 38 |
| 11 | | 46 | 52 |
| Averages | | 31 | 45 |
| P. E. of Av. | | 3.02 | 1.90 |
| Group C | | | |
| Pupil | | | |
| 1 | | | 58 |
| 2 | | | 46 |
| 3 | | | 34 |
| 4 | | | 39 |
| 5 | | | 28 |
| 6 | | | 22 |
| 7 | | | 20 |
| Average | | | 31 |
| P. E. of Average | | | 3.7 |

knowledge of this material is larger than in the case of "The Admiralty Islanders." The effect of reminiscence which is noted here may of course be merely due to the fact that the pupils have looked up some of the answers to the questions or it may be due to repeated practice with the test.

One may take the averages above and use them as corrections for the various uncontrolled factors and arrive at the following:

| | |
|---|---|
| Effect of previous knowledge | 28% |
| Effect of motivation | 15% |
| Effect of single reading | 9% |
| Effect of immediate recall on delayed recall | 8% |

The above corrections are likely to be in error either plus or minus to a considerable degree. The reader is probably justified, however, in drawing some conclusions in regard to the approximate effect of the various factors in a single reading of this type of material. There cannot help but be a large factor of previous knowledge in this selection since it is common material found in the ordinary language book. In general, we can say that the single reading affected the scores on immediate recall to a very slight extent. This is quite in keeping with expectation, since the material is hard and full of detail. Introspection would tell one that this type of material is not retained to any large extent after a single reading.

### 7. Conclusions from this Chapter

The general conclusions from this chapter are as follows:

From the experiment with "The Admiralty Islanders" the following conclusions are drawn:

a. The effect of previous knowledge on the scores is large in comparison with the total gross scores.

b. The effect of a single reading when corrected for previous knowledge is comparatively small; it doubles the knowledge, but it still leaves the pupil far short of perfect mastery of the ideas in the selection. The amount actually added to previous knowledge ranges from 26 per cent in grade 8 to 6 per cent in grade 4.

c. The effect of the use of an initial test is great. It ranges from 14 per cent in the eighth grade to 7 per cent in the fourth grade. That is to say, on the average 14 per cent of the total ideas of the selections are added in the eighth grade because of the use of the initial test, or because of that and inequality in the groups. It seems probable that giving an initial test before a reading of any article which the pupils are assigned is a good method of insuring mastery.

d. The effect of an immediate recall test upon a delayed recall is practically the same as the motivation caused by the initial test. To be more explicit, the effect of the review preceding a delayed recall is very marked. This is in accordance with our knowledge of the laws of memory.

e. If a group of children read a selection and have no immediate recall of the facts, such a reading does not add materially to the knowledge which they already have. This would seem to show that frequent tests are a good device as a means of increasing ability to remember. This, too, bears out the truth of the law of recency in memory.

f. The effect of a single reading and two tests before a delayed recall is decidedly marked and is distinctly greater than the effect of one test and a single reading. This would seem to show that the greater number of repetitions of the material, even in the form of a test, the greater the chance for remembering the material at a later time.

g. For this type of material the general conclusion may be made that for the ordinary child a single reading is not an efficient device. There are, however, some individuals who are so skillful that the single reading for them is relatively sufficient. Perhaps it would be clearer to say that the efficiency of a single reading varies with the individual in the same manner as other abilities which have been measured. There are some children for whom it is without appreciable effect.

h. The series of corrections established in connection with this experiment may be used only tentatively to estimate the probable effect of various factors on other materials. The implications are very suggestive, however, in the absence of anything more accurate.

i. The experimenter does not mean to imply that for cursory reading of certain types of material the single reading is not a desirable device. Newspaper reading, reading for pleasure, and the like need consume no more time than a single contact will give; but for efficiency in study of factual material it is patent that a single contact will not do.

The following conclusions in regard to the experiment with the material "On the Use of Abbreviatians" are added to those given in the discussion following the tables.

a. The material was much more familiar to the pupils than the material "The Admiralty Islanders" as was to be expected.

b. The same general results are apparent in the tables.

c. The chief difference is in the apparent reminiscence in the delayed recall which may be due to the shortness of the interval between immediate and delayed recall or to the fact that the pupils may have looked up the answers in the interval between the two tests, or to actual reminiscence.

d. This experiment reveals that on very factual material of a difficult character, the amount of effect of a single reading is so small as to be practically *nil*. This is material for hard study and drill. No teacher would be likely to ask a pupil to study it in the ordinary way; but if she did, the results would soon put her right in regard to the inadequacy of such a procedure.

# CHAPTER VII

# THE EFFECT OF A SINGLE READING AND THE EFFECT OF TWO READINGS COMPARED

The foregoing chapter attempts to get at the effect of variable factors which enter into a single reading. It is significant that previous knowledge, which no doubt is present in all reading since it furnishes a part of the apperceptive mass which makes it possible for a reader to become interested in and read a selection, is considerable in amount. It cannot be measured except by means of a test before reading and even then it is difficult to know whether or not the test represents previous knowledge entirely or a complex made up of previous knowledge of the material, natural ability to reason out the answers to the questions, and chance checking of the right answers. The effect of the metal set caused by giving this test is also mixed up in the results which are obtained by testing after a single reading preceded by a test. Practically, the teacher and administrator cannot give much consideration to these factors, except to know that they are present and to appreciate that the result of a single contact with material is not great unless that contact is highly purposeful. Therefore, in ordinary testing of school reading, these factors are disregarded and the results obtained after one or more readings of the article are attributed to that reading. In this chapter, some significant material is reported which is obtained by testing equated groups in the ordinary fashion. The method of equating the groups in this series of experiments gives them some significance.

This chapter will throw some light on the effect of a single reading by comparing it with the effect of an immediately repeated reading of the same kind. Just as in the field of physical measurement an inch is an arbitrary unit and has no absolute meaning except in comparison with some other unit of measurement, so in measuring mental products, perhaps, the effect of a single reading is more adequately expressed by comparing it with what happens when the same article is read over again. The following is an account of an experiment with 228 pupils in the seventh and eighth grades of a typical middle-western departmentalized intermediate school.

### 1. PURPOSE

The purpose of the present experiment, while generally stated above, is more particularly as follows:

a. To obtain a standard of comparison which may allow one to say that the effect of a single reading is a half as large as the effect of two readings, or twice as large, or whatever the case may be.

b. To ascertain if it pays to say to pupils, "Read your lesson twice, for this is an economical method of study."

### 2. MATERIAL AND METHOD

The material used in this experiment was the article "Medieval Castles," the construction of which was explained in a previous chapter.

In order to get a measure of the effect of two readings, the pupils in each class were divided into two groups by ranking them on their comprehension scores in the Monroe Silent Reading Test III, and then putting number 1 in one group, number 2 in the second group, number 3 in the second, and 4 in the first, etc. In this way two approximately equal groups were formed in each class. Group A read the selection a single time; Group B read the selection twice. Both groups were then given the test to determine their comprehension and remembrance of the article read. Twenty days later the same test was repeated in order to measure the effect of lapse of time upon memory of this material and to compare the effect of one reading and two readings as measured by delayed recall.

The approximate equality of the groups is shown by the following gross Monroe Scores for each class:

| Room | No. of Pupils | Group A | Group B |
|---|---|---|---|
| Acceleration Room | 14 | 249 | 246 |
| West 8 A | 24 | 379 | 384 |
| East 8 A | 18 | 264 | 266 |
| 8 B | 22 | 284 | 291 |
| West 7 A | 32 | 413 | 412 |
| East 7 A | 18 | 185 | 183 |
| 7 B | 28 | No score | No score |
| Total | 156 | 1774 | 1802 |

The Monroe Comprehension score was used for division into groups because the ability to do the comprehension tests of Monroe most nearly resembles the ability to perform the tests in the present experiment. Rate of work was not considered because this is a test of amount of difficult material that could be done and not of time taken to perform the tests; moreover it is doubtful if the two variables can be measured at the same time (30). Pupils had all the time they needed to perform the tests.

The directions given to the pupils were the same as in former experiments with the exception that Group B was told to read the article twice. Ample time was taken to explain to the pupils the procedure in the experiment. While reading the material, pupils were asked to mark every half minute by putting a line on the reading page, similar to the procedure in the Courtis Silent Reading Test No. 2. This was for the purpose of getting some check on whether or not a single reading took place with Group A and whether or not Group B read the selection twice. Very little was said about the speed with which the article was to be read except, "Read in your usual manner". Time was kept on both readings and the test by the method explained in the previous experiments. Rates of reading are reported on this material because of the interesting light they throw upon the probable manner in which the two groups read. It is believed that though no other directions in regard to rate of reading than above were given, some children may have hit on the plan of reading carefully once and skimming a second time, or the reverse. The rate scores will indicate if this was done.

### 3. RESULTS

The results of the experiment are given in the following summary tables. Pupils who were not present for both tests—immediate and delayed—are excluded. The total number of pupils remaining after those were eliminated because of irregularities is 156 pupils or 78 pairs in grades seven and eight, and in the Acceleration Class made up of pupils approximately eighth grade in ability.

TABLE 37

EFFECT OF A SINGLE READING AS COMPARED WITH THE EFFECT OF TWO READINGS

*Immediate Recall*

| Grades | GROUP A<br>One Reading | | GROUP B<br>Two Readings | |
|---|---|---|---|---|
| | VIII | VII | VIII | VII |
| *Class Interval* | | | | |
| 70 – 74.9 | | | 1 | |
| 65 – 69.9 | | | 1 | |
| 60 – 64.9 | 1 | | 2 | 2 |
| 55 – 59.9 | 5 | | | |
| 50 – 54.9 | 4 | | 6 | |
| 45 – 49.9 | 3 | | 4 | |
| 40 – 44.9 | 3 | 3 | 4 | 2 |
| 35 – 39.9 | 6 | 5 | 4 | 10 |
| 30 – 34.9 | 5 | 2 | 3 | 2 |
| 25 – 29.9 | 5 | 8 | 5 | 7 |
| 20 – 24.9 | 3 | 8 | 6 | 8 |
| 15 – 19.9 | 2 | 5 | 2 | 1 |
| 10 – 14.9 | 2 | 5 | 1 | 4 |
| 5 – 9.9 | | 2 | | 2 |
| 0 – 4.9 | | 1 | | 1 |
| Total Cases | 39 | 39 | 39 | 39 |
| Median | 37 | 24 | 38.1 | 27.5 |
| Q | 11.2 | 7.2 | 12.3 | 8.1 |

The above table should be read as follows: on a single reading, one pupil in group A made a score of from 60 – 64.9 per cent of the total possible score, five pupils a score from 55 – 59.9 per cent of the total possible score, etc. The median score from the eighth grade on a single reading is 37; on two readings, 38.1 per cent or 1.1 per cent better. The score for Group A, grades seven, for a single reading is 24 per cent of the total possible score; for Group B which read twice, 27.5 per cent of the total possible score, or 3.5 per cent better than the single reading group. Thus, the medians for the groups which read the selection twice are slightly better than those which read it once. The quartile deviations, Q, show that there was a slightly larger spread after two readings than after a single reading.

### *Conclusions from Preceding Table*

1) The effect of two readings is very little if any greater than the effect of a single reading of this material for the pupils represented in this experiment as measured by the median scores made by them.

2) The only appreciable sign of greater efficiency due to two readings is the slightly greater spread of the Group B which read the article twice. This may be due, however, to the fact that the groups were not accurately divided in spite of the care taken.

3) If these scores are to be taken as a sign of efficient reading, there is still considerable to be done in the matter of advising pupils in regard to economy of study. Apparently two readings

of the article are not efficient from the standpoint of immediate recall. These pupils would have been practically as well off if they had read the article a single time and then gone on to something else.

4) A study of the answers made to these tests with a determination of whether or not the two readings resulted in a selection of the more important ideas would help to determine whether there was any gain made in the reading which is covered up by the median group scores.

5) A single reading must be followed by something other than a mere repetition of the same kind of reading, if an increased number of ideas are to be reproduced.

6) These data show that the largest gain from reading is due to the single first reading; if this is true then training in a single careful reading so that the average pupil can reproduce seventy-five to ninety per cent of the ideas in the article would seem to be the first step in economy of study.

7) If, as Dr. Germane found (23), rereadings are more efficient than a corrected summary, and the directed summary more efficient than undirected rereadings, then we seem to be on the trace of some very valuable suggestions in regard to efficient methods of study.

The following table gives a summary of the scores made on delayed recall by the same groups.

TABLE 38

EFFECT OF A SINGLE READING AS COMPARED WITH EFFECT OF TWO READINGS AS MEASURED BY DELAYED RECALL

| Grades | GROUP A Silent Reading | | GROUP B Two Readings | |
|---|---|---|---|---|
| | VIII | VII | VIII | VII |
| *Class Interval* | | | | |
| 70 – 74.9 | | | | |
| 65 – 69.9 | | | | |
| 60 – 64.9 | | | | |
| 55 – 59.9 | | | | |
| 50 – 54.9 | | | | |
| 45 – 49.9 | | | 1 | |
| 40 – 44.9 | 3 | | 4 | 1 |
| 35 – 39.9 | 9 | | 5 | 1 |
| 30 – 34.9 | 1 | 1 | 2 | 1 |
| 25 – 29.9 | 5 | 1 | 2 | 4 |
| 20 – 24.9 | 7 | 3 | 5 | 4 |

| | | | | |
|---|---|---|---|---|
| 15 – 19.9 | 5 | 13 | 10 | 9 |
| 10 – 14.9 | 6 | 6 | 6 | 6 |
| 5 – 9.9 | 2 | 12 | 4 | 10 |
| 0 – 4.9 | 1 | 3 | | 3 |
| Total | 39 | 39 | 39 | 39 |
| Median | 23.9 | 13.7 | 19.7 | 15.2 |
| Q | 9.97 | 5.18 | 10.28 | 9.10 |

The above table is read in just the same manner as Table 37. The median for the group which read the article once in the eighth grades is 23.9 per cent of the total possible ideas; for Group B which read the article twice, 19.7 per cent of the total ideas or 4.9 per cent less than that of the group which read only once.

The median for Group A (single reading) in grade seven, is 13.9 per cent of the total possible ideas; of Group B (two readings) it is 15.2 per cent or 1.3 per cent better.

The quartile deviation, Q. is greater for the groups which read the article twice.

### *Conclusions from the preceding table*

1) As far as the median scores for the groups show, the pupils who read the article once only in the eighth grade were better off on the average than those who read it twice.

2) In the seventh grade, the pupils who read the article twice were more efficient in reproducing than the single reading group, but by only a slight margin.

3) As measured by delayed recall, the loss due to lapse of 20 days is as follows:

| | Group A | Group B |
|---|---|---|
| Grade 8 | 13.6 | 11.4 |
| Grade 7 | 10.3 | 12.3 |

Group B thus lost more through lapse of time than Group A. This may possibly be due to the fact that on immediate recall they remembered more of the details than Group A; but that on delayed recall these details tended to be forgotten.

4) For some unaccountable reason, Group A in the eighth grade did better than Group B. This may be due to the fact that it is actually a better group of readers, or it may be due to the fact that on this kind of material, a second reading is of slight advantage, and for delayed recall may be a disadvantage for it

may center the attention on the details rather than the main ideas which is not of advantage for delayed recall.

5) From these data one is led to suspect that the effect of repetition on logical material of this type, as on nonsense material, does not begin to be appreciable until after many such repetitions have taken place. For the ordinary reader probably the best general advice is, "Read the article through carefully once." The effect of summarizing, as pointed out by Dr. Germane, depends upon the kind of summary used. Directed reading is much more efficient than even repeated undirected reading.

### *A Study of Scores by Paired Individuals*

It is a well-known principle that gross scores often cover up the truth about the individuals whose scores make up the averages or medians. In this study a precaution has been taken to insure getting at the whole truth if it is possible in group measuring. The following pages contain a study of the effect of a single reading as compared with the effect of two readings from the standpoint of the number of individuals in the single-reading group (Group A) who excel or are excelled by their controls in the two-reading group (Group B). Table 39 gives these data for immediate recall; Table 40 for delayed recall. The pairs are as nearly equal in reading ability as it is possible for one to get them by means of Monroe Silent Reading Scores, Test III.

TABLE 39

EFFECT OF ONE READING COMPARED WITH EFFECT OF TWO READINGS AS REVEALED BY STUDY OF INDIVIDUALS AND THEIR CONTROLS

| Grades | No. A's Excelling B's | No. B's Excelling A's | Tied | Adv. for B's | Ave. % Adv. |
|---|---|---|---|---|---|
| 8- Accel | 0 | 7 | | 7 | 10.5 |
| W 8 A | 7 | 5 | 1 | –2 | .5 |
| E 8 A | 6 | 2 | 2 | –4 | –8.1 |
| 8 B | 5 | 5 | 1 | 0 | 3.9 |
| W 7 A | 5 | 11 | | 6 | 12.5 |
| E 7 A | 3 | 6 | | 3 | 3.6 |
| 7 B | 5 | 8 | 1 | 3 | 4.1 |
| | 31 | 44 | 6 | 13 | 6.2 |

The above table should be read as follows: Of the 7 pairs of pupils in Grade 8, Acceleration Room, 7 individuals in Group B, the two-reading group, excelled their controls in Group A, the one-reading group, etc. Column five shows the superiority of Group B, the two-reading group, as expressed in number of individuals who excelled their controls; column six expresses the average per cent of reproduction by which Group B excelled Group A. A minus sign indicates that Group A excelled Group B.

Table 39 shows the following:

1) On immediate recall 31 individuals in Group A, the single-reading group, excel their controls in Group B, the two-reading group, but 44 pupils in Group B excel their controls. This gives Group B an advantage of 13 individuals.

2) In all the grades but two Groups B, the two-reading groups, excel Groups A who read once.

3) The average amount by which Group B excels Group A in all grades is 6.2 per cent of the total ideas contained in the article.

4) This shows a slight superiority for Group B which read the article twice, but the amount of superiority is not great. The second reading adds only a slight increment to the first reading, if these data are to be considered reliable.

A similar presentation of data for delayed recall is found in Table 40, which is read the same as Table 39.

TABLE 40

EFFECT OF ONE READING COMPARED WITH EFFECT OF TWO READINGS AS REVEALED BY STUDY OF THE INDIVIDUALS IN EACH GROUP WITH THEIR CONTROLS

| | | Delayed Recall | | | |
|---|---|---|---|---|---|
| Grade | No. A's Excelling B's | No. B's Excelling A's | Tied | Adv. for B's | Ave. % Adv. |
| 8 Accel. | 0 | 6 | 1 | 6 | 6. |
| W 8 A | 6 | 5 | 1 | –1 | –6.9 |
| E 8 A | 4 | 3 | 3 | –1 | –3.9 |
| 8 B | 3 | 6 | 2 | 3 | 3.8 |
| W 7 A | 7 | 9 | 0 | 2 | 3.7 |
| E 7 A | 2 | 7 | 0 | 5 | 3.6 |
| 7 B | 8 | 5 | 1 | –3 | – .7 |
| | 30 | 41 | 8 | 11 | 8. |

The foregoing table seems to show that:

1) On delayed recall the advantage remains with Group B, the two-reading group, which excels Group A both in the number of individuals who excel their controls and in average per cent reproduced.

2) In all grades but three, Groups B excel Groups A.

3) The average by which the individuals in Groups B excel Group A is 8 per cent of the total possible score.

4) This shows a slight superiority for Group B and suggests that the effect of two readings on delayed recall is slightly better than on immediate recall.

### *Conclusions*

When measured by the careful comparison of individuals in Group A with individuals in Group B, grade by grade, the following conclusions concerning the effect of two readings, as compared with the effect of a single reading, seem to be allowable:

1) There are certain advantages in the second reading which are covered up when the scores are handled in a distribution table. The difference between the Groups is more marked.

2) On immediate recall there are thirteen more individuals in Group B who excel their controls in Group A then the reverse; the average amount, while small, is significant.

3) The fact that on delayed recall eleven more individuals in Group B excel their controls than the reverse by an average amount of 8 per cent seems to indicate that a repetition of the reading, while small, is measurable by group testing.

4) It seems probable that in reading logical material in this fashion, as in learning nonsense syllables, many repetitions would be necessary before the results of the reading would be practical mastery. It is doubtful, however, whether many succeeding readings would be worth while. Probably directed reading, summarizing, or some other device would be much better than a number of rereadings of this character.

5) These data show that the first reading, while it results in the mastery of only a small part of the facts in material of this character, results in more knowledge of the material by far than is added by a second reading similar to the first.

6) The largest amount of gain on the average for a second reading is a possible superiority of 6 per cent on immediate recall and 8 per cent on delayed recall.

### *Effect of one and two readings as measured by rate*

It is probable that if a pupil knows that he has time to read an article twice his rate of reading may be quite different from that when he has only a short time to read the material and knows that he can read it only once. In this section data are given to show the effects of a single reading and of two readings as measured by rate. The purpose is to show:

1) Whether Group B reads faster or slower than Group A.

2) Whether or not Group B reads faster the second time than the first.

3) What the probable effect of knowing that they can read twice has upon the rate of Group B.

The reading rates given herein are not exact to the second but represent approximate rates which were obtained in the following manner. The experimenter wrote on the blackboard a series of numbers at ten second intervals. When a pupil finished his reading he wrote down on his paper the number then appearing on the blackboard. If he finished in the interval between 50 and 60 seconds, he marked 5 on his paper and this figure was used in computing his rate. The rates of reading were found by dividing the total number of words into the total number of seconds read and then figuring the approximate number of words read per minute. Table 41 gives these rates for both groups.

TABLE 41

COMPARISON OF READING RATES OF GROUP A AND GROUP B ON ONE AND TWO READINGS

| Grade | No. Prs. | Group A | Group B | |
|---|---|---|---|---|
| 8 Accel. | 7 | 176 | 180 | 202 |
| W 8 A | 12 | 165 | 182 | 189 |
| E 8 A | 9 | 169 | 152 | 160 |
| 8 B | 11 | 146 | 144 | 126 |
| W 7 A | 16 | 167 | 167 | 174 |
| E 7A | 9 | 140 | 126 | 126 |
| 7 B | 14 | 133 | 171 | 155 |

The rates given are average rates for the grade and group. The first column under Group B indicates the first reading; the second column, the second reading.

A study of the table will reveal the following:

1) In four grades out of seven, Group B read faster or equalled the rate of Group A on the first reading.

2) In all grades except two in rereading Group B read faster or equalled its rate on the first reading.

3) In four grades out of seven Group B read faster on second reading than on the first reading.

### *Conclusions on Rate of Reading*

1) Group B in the majority of cases read faster during the second reading than the first. This suggests that the first reading was a careful reading and the second a faster reading for the purpose of verifying or checking up.

2) Group B consistently read faster than Group A. This may be due to the fact that they knew that they could reread so did not think it necessary to read carefully the first time.

3) Although it may be possible that Group A naturally reads slower than Group B, the known high correlation between speed and comprehension seems to make it hardly probable, since the groups are so nearly equal in reading ability as measured by this test.

4) Had Group B been told to read slowly and carefully both times the effect upon the ability to do the test would probably have been to increase the amount reproduced.

5) The probable effect of the knowledge that they may reread seems to have resulted in higher rates for the rereading group and may have slightly lowered the rate of the single reading group on account of the fact that they knew they were to be tested.

6) These data suggest many interesting problems such as: "What is the effect of different rates upon the ability to reproduce?" and "What would have been the effect had Group A been told to read slowly and carefully and Group B told to read through swiftly both times?"

### 4. General Conclusions Derived from This Chapter

The effect of a single reading compared with the effect of two readings of approximately the same character would seem to minimize the effect of unmotivated repetition under ordinary schoolroom conditions such as were represented here. The second reading added only 6.1 per cent on immediate recall and only 8 per cent on delayed recall. This would suggest that the necessity for other methods of study is great. Some individuals might hit on the device of reading for the gist the first time and rereading for details the second and thus profit by the practice but it is hardly probable that elementary school pupils would do this unassisted by suggestion from the teacher. Rather than to tell pupils to reread it seems a much wiser practice to direct their reading by a series of questions, or to train them to a high degree of skill in reading an article a single time and mastering the salient facts. While a single reading ordinarily results in a mastery of only a small per cent of the ideas in an article of this character, rereading adds only a small amount to that obtained from a single reading.

# CHAPTER VIII

## RELIABILITY OF THE STUDY

It seems desirable before concluding this study to discuss in a single chapter the reliability of the materials used, the method, and the conclusions. Whatever evidence may be at hand in regard to the reliability of doubtful points in this study is largely the work of the experimenter, but here and there he has accepted the conclusions of other experimenters as a point of departure. Wherever it has been possible to secure experimental evidence in regard to a point, this evidence has been used, but in some cases the experimenter has proceeded in the best method known at the present time without subjecting the method to experimental testing. He regarded it more important at this time to try out several types of materials and several types of tests than to concentrate on any one narrow type. Further experimentation should be done to get added evidence where there are doubtful points.

### 1. The Materials.

#### a. *Reading Materials*

The reading materials used in this series of experiments are of varying lengths. They are interesting, highly factual, and representative of the kinds of materials ordinarily found in the elementary course of study. The writer did not find as did Gray (14) that long materials tended to fatigue pupils and that they did not do as well toward the end of the test as toward the beginning, if the test were of the completion or true and false type. It is true, however, of the type of test which makes large demands on recall. Since the writer used the completion and the Yes and No test in the more important experiments, he feels that the length of the articles did not result in a loss toward the end. Indeed, he had evidence in a carefully analyzed group of papers from 26 pupils that such was not the case, but that pupils held up until the end. This table is too long to include here.

There is another just criticism that may be brought against the materials and that is the manner in which they are printed. Huey and Dearborn found that the ideal book line was about four and one-quarter inches in length. The writer fails to find, however, that in experiments with six different types of material on the same group of 16 pupils there was any difference in the reading rates of these pupils even when the length of lines and the character of the printing differed widely. It is the character of the reading material, rather than the kind of printing that slows the rate, that is, below a certain point, and the eye strain on so short a selection cannot be great. Moreover, a typewriter with large type was used for printing reading materials which were to be hectographed and this tended to overcome any disadvantages to which pupils may have been put.

There is very little criticism that can be offered in regard to the kinds of materials used since they were of an interesting, factual character specially suited to silent reading experiment. Some of the materials were doubtless very difficult for the lower grades but even in grade four many pupils made very good scores as a result of a single reading.

### b. *The Tests*

One of the most moot points in the testing of silent reading is in regard to the type of test employed. Criticisms are made of all the tests which have so far been used. The tests which have been most favorably received, however, are of the same general type as those used by the writer—tests requiring little recall and depending chiefly upon recognition. The silent reading tests of Courtis, Monroe, and Thorndike are of this type. The completion test is a better measure of the effect of reading because it furnishes a better stimulus to recall than a pure question and answer test, or a reproduction test requiring the child to compose an answer in writing. For these reasons, in these experiments where the object was to measure associations which had been formed but were not strong enough for recall without stimulus, the completion and the true and false tests were the better type to use.

The next question to raise is whether or not the tests actually did test silent reading. This is a question which must be answered by all the various experimenters in the field. The fact that the tests show a decided effect from the result of a single contact with the reading material, and that when there was no reading a repetition of the test did not materially raise the second score, tends to show that the test is a test of reading. The following summary of an experiment made with 22 seventh grade pupils to determine this point gives evidence to the effect that the test is a measure of reading.

The completion test for the material "Peanuts" was given to 22 pupils with two minutes and a half intervening with the following results: the average amount reproduced with the first test was 22 per cent and with the second test 21.5 per cent. Pupils who had read the article, on the other hand, made the following scores:

| Grade | VIII | VII | IV |
|---|---|---|---|
| After reading | 27.2 | 36.2 | 18.9 |
| Before reading | 21.9 | 15. | 10. |

Since the reading was the only uncontrolled variable in the situation, it must have been the effect which was measured by the test. As to just how rigidly this test brought out the effect of reading is another question. These experiments give evidence to show that the completion test and the true and false test certainly result in larger scores than the question and answer test.

Another evidence of the reliability of the tests used is found in the nature of the scores which resulted. When it is found that the scores exhibit the same characteristics as any other measurement of mental ability—wide individual differences, grouping about a central tendency, and rather consistent variability from grade to grade, with a distinct gradation of ability on the average and much overlapping among the grades, it is quite evident that the tests have been a more or less reliable measure of some ability possessed by the pupils. All of these characteristics are consistently exhibited in the results of the tests used by the writer. As to their relative reliability, that is a question which may be decided by correlating the scores of pupils in these tests with other measures of reading ability. The writer is not able to furnish those correlations for the pupils in three schools owing to the lack of other measures of reading ability. And while the complete correlations have not been worked out for the fourth school, the writer knows from a preliminary examination of the scores made by pupils in the tests of these experiments and other silent reading tests that the correlation is high with the single exception of the Kansas Silent Reading Test. Examination of the measures of variability in connection with the tables in the main chapters of this study will furnish evidence as to the consistent variability from grade to grade. These facts and a comparison of the scores made by the pupils in these experiments with other silent reading scores, will sufficiently establish the reliability of the tests.

There is some question as to whether the completion test is a measure of reading ability or not. The uncontrolled completion test is much used as a measure of general intelligence. Trabue used it to test language ability. It is probably true that any test of reading is to a large extent a test of general intelligence. The type of test used in this study, however, is a controlled test in which there are five words suggested as the word needed to complete the true meaning of the statement. The fact that pupils consistently improved in their answers to the test after a single reading shows that the test did measure the effect of the only uncontrolled variable in the situation—the single reading. And the fact that the association was suggested and did not depend upon the ability to supply the right word from memory but rather upon the ability to recognize the truth of the statement when the right word was selected, would eliminate to a large extent the effect of superior language ability on the part of some students. It is recognized, however, that pupils of superior intelligence do better than those of inferior intelligence in this test. But the pupil of superior intelligence also reads better. For these reasons the writer cannot see how the question of intelligence in general is to be separated from the reading complex. The ability to read implies intelligence, and a test of reading is a test of general intelligence.

The true and false test used in one of the later experiments of this series is open to criticism on the ground that there is a large factor of chance answers. The writer believes those answers which he checks correctly in both positive and negative forms are an actual measurement of his knowledge. On the other hand, it may be that there is an advantage in stating the question in one form and that the pupil may actually know the fact in question but be confused by the form of the statement and checks the wrong answer. The following table will give the differences between the total score of three groups as measured by the total number of right answers checked and the total number of paired answers.

TABLE 42

CORRECT ANSWERS ON "ABBREVIATIONS" GRADE VIII AS MEASURED BY BOTH SINGLE AND PAIRED ANSWERS

| Group | Initial Test Single | Initial Test Pairs | | Immediate recall Single | Immediate recall Pairs | | Delayed recall Single | Delayed recall Pairs | |
|---|---|---|---|---|---|---|---|---|---|
| A | 51 | 28 | 23 | 65 | 46 | 19 | 66 | 47 | 19 |
| B | | | | 62 | 31 | 31 | 62 | 45 | 17 |
| C | | | | — | | | 56 | 31 | 25 |

It is seen from the above table that the number of correct answers as measured by the single answers correctly checked is sometimes twice as large as the number when measured by the paired answers. By subtracting the number under the column "Pairs" from the column "Single" we find that in each case the answers checked right according to the single answers, exceed the paired answers somewhat the same in amount, ranging from 17 to 31. These answers were checked right by chance, or else some of them were suggestive in the positive or the negative form and elicited right answers. The writer demonstrated that this factor would operate by checking at random alternate Yes and No answers in several different ways without reading the statements at all, with a result that from 41 per cent to 52 per cent of correct answers were correct by chance. It is very doubtful, however, that a child would check in this way. In fact it is impossible to estimate the exact amount of knowledge the child possesses with this type of test; by discounting the answers one may be ruling out answers that are the actual result of the child's knowledge of the facts. An adult might hit upon the scheme of chance checking but it is unlikely that the child would.

Nevertheless, in order to eliminate this effect of chance answers, the writer instructed pupils carefully in the main experiments not to guess at the answers. Then to make the matter doubly sure, he constructed the true and false test used in the experiment with the material "On Abbreviations" so as to eliminate the effect of chance. There are 50 paired answers in this test, or 100 points in all. A question must be answered in both positive and negative form before it is counted as correct. The results of checking the single answers against the paired answers shows that the child makes a much larger score on single answers than he does on the paired answers. This may be taken to mean that safest assumption is that they were checked correctly by pure chance. The consistency with which pupils marked pairs of answers correctly would argue against the assumption that the statement of the fact in two forms was confusing and resulted in many pupils becoming confused and changing their minds after having checked an answer rightly the first time, thus making a mistake in the second. Taking all these facts into consideration, one may say that while there is a possibility of this type of test failing to measure completely the effect of reading it probably measures comprehension as well as any type of test which has been devised. It has the advantage of being very objective, easy to do, and it furnishes a rigid check on comprehension. The fact that pupils on delayed recall reproduced better than on immediate recall argues that the test did not present great difficulty as a test. This is further substantiated by the fact that the papers were quite consistently complete and were as consistently filled out at the end of the test as at the beginning. Yet in spite of this care to eliminate chance, the per cent of correct answers reproduced does not differ materially from those gained by the use of the controlled completion test in Chapters VI and VII.

Finally, the writer would call attention to the fact that the tests measure the effect of reading in a much more exhaustive way than ordinary tests because the pupil had all the time he desired to complete the exercises. The chances are that a very complete measure of the effect of reading is shown in the various tables that comprise this study.

## 2. The Method of Experimentation

The method of experimentation used in this series follows out the recommendations of former experimenters and writers on experimental method in education. Care was taken at all times to put pupils at their ease, to explain fully and carefully until all understood what was desired, and to keep the conditions of the experiments, which were subject to control, as constant as possible. The time of day was not regulated but it is now quite generally believed that the time of day affects very little a pupil's performance in an experiment of this kind. Conditions of a physical nature are of less importance than the pupil's mental attitude toward the work. For this reason experiments were given whenever pupils could be conveniently obtained for them. If there is any advantage in time of day, so many different times are represented in these experiment that variations one way are offset by those in another.

The method of selection of the groups for the experiments in the sixth chapter is open to some criticism. It is probably doubtful whether or not in the three groups selected each represents a very close approximation to the normal group in the grade from which they are taken in schools where the division was made by teachers' judgment. However, in two other schools where pupils were divided on reading test scores, the scores were roughly comparable and in general approximately the same as in the schools where division was made by teachers' judgment. They were therefore all included in the final tabulations. It is believed that these scores are sufficient in number to be significant and that the correction made as a remedy of these scores would not be greatly changed by additional testing. The length of time required for these experiments also made extension of the tests possible. It is a well recognized fact, moreover, that we can sample at any time only a small part of the total school population; all standards are made on the basis of sampling of the total abilities represented in the school population. When we have, as in this study, particularly in the experiments with the material "Peanuts," "The Admiralty Islanders," and "On the Use of Abbreviations," a selection of pupils from the same grades in several different schools, a considerable dependence can be placed upon the results. Wherever the cases are few and the results tentative the experimenter has pointed out that fact.

## 3. Scoring the Tests

Before undertaking the scoring of the present study the writer considered several methods of marking the papers. He accepted the method used by former experimenters in similar studies and decided to mark one point for each correct answer and to give nothing for an incorrect answer. Now obviously, all answers do not represent the same amount of work; not all questions are of equal difficulty. If a pupil answers a question which is missed by every other member of a group he should receive more credit for that answer than for an answer which is made by every other member of the group. This is on the theory that every member of the group has had the same chance to answer a question and that some of them have read the question and considered it too hard to answer. To overcome this inequality in difficulty the common plan has been to resort to scaling the questions and giving to each question the value corresponding to its difficulty as shown by the number of pupils taking the test who have missed the question. This is doubtless the most scientific method of marking the results of the tests and where great accuracy is desired it should be used. However, on certain types of material the effect of marking by this system as contrasted with the point system of marking used in this study, has been greatly overestimated. The writer undertook a study to show how scaling the answers would affect the scores in connection with the present study. He marked a series of papers by both methods and then correlated the answers of a group of sixteen pupils. The results of this study are shown in the following table. Where a very fine measure of the differences in rank between students is counted and where extensive correlations are to be run, the method of scaling the answers is more desirable. In the present study, however, where there is no fine distinction to be drawn between individuals the point system used is quite accurate and gives results which justify the general conclusions that are drawn. This opinion is supported by evidence given by Pressey and Pressey (31), in an article in *School and Society*, June, 1920. They conclude that "Monroe's weighting of

TABLE 43

MARKS OF PUPILS ACCORDING TO EQUAL AND WEIGHTED VALUES OF QUESTIONS

*Medieval Castles*

| Pupil | Marks by Unweighted Method | Marks by Weighted Method |
|---|---|---|
| 1 | 64.3 | 52.8 |
| 2 | 57.1 | 38.6 |
| 3 | 43. | 34.5 |
| 4 | 29. | 16.8 |
| 5 | 21.7 | 6.14 |
| 6 | 14.2 | 5.8 |
| 7 | 35.7 | 24.9 |
| 8 | 30.9 | 22.9 |
| 9 | 28.5 | 18.5 |
| 10 | 59.5 | 53.9 |
| 11 | 61.9 | 50.4 |
| 12 | 11.9 | 7.6 |
| 13 | 21. | 16.8 |
| 14 | 30.9 | 24.7 |
| 15 | 11.9 | 3.7 |
| 16 | 45.2 | 35.3 |
| Averages | 35.5 | 24.4 |

his comprehension score adds practically nothing to a crude score as a measure of reading ability.''

It is seen from the above table that the effect of using a weighted system of marking is to reduce the range of marks about 10 per cent. It is to be remembered, however, that several questions by the second method are counted as if of no value because of the fact that they were answered by all the pupils. The effect of this is to reduce the range of answers to some extent.

The correlation between the two systems of marking according to Pearson's Product-Moments method, was found by the writer to be .89. This is a high correlation and shows that the results would not have been markedly different in character had the weighted system of marking been used.

Moreover, in this system of marking the effort was made to express as a per cent the number of ideas reproduced as related to the total number of ideas in the article. No attempt was made to evaluate these ideas, to mark them as difficult or easy, or in any other way determine the difficulty of the questions. It was thought sufficient for the purposes of this study to show the effect of a single reading as measured by the average amounts reproduced of the whole number of ideas in the selection. The weighted method of marking would have had some disadvantage in that it does not represent the per cent of the ideas reproduced as the result of a single reading but rather the efficiency of the reading as measurd by the weighted values of questions. The writer does not believe that there would have been any advantage in using the weighted method in this study.

# CHAPTER IX

## SUMMARY AND GENERAL CONCLUSIONS

A general summary of the facts gathered in the present study and a comparison of these facts with those gathered in previous studies which were reviewed in Chapter III will be given here. This summary will be followed by the general conclusions of the present study. Table 44 presents in a tabular form the scores made on various types of material by pupils examined in the present series of experiments. Table 45 makes some comparisons between the results of previous studies and those of the present investigation. It may be pointed out here that these data are practically all the significant data bearing on the problem, "What is the effect of a single reading?" Facts of considerable significance which emphasize the inadequacy of such a direction as "Read your lesson as many times as you can" are here presented. Former experimenters have apparently given little or no attention to this problem. The materials included in the foregoing chapters, while not free from error and often tentative, bring out the importance of the work and contribute to a solution of the problem set in the opening chapter. They suggest that the complicated problem involved in reading is only on the way to being solved and that there are matters which have been merely touched upon that yet offer opportunity to future investigators for studies which will add incomparably to our knowledge of reading and study.

TABLE 44

SUMMARY OF FINDINGS OF PRESENT STUDY

| Article | Length | Character | No. Pupils | IV | V | VI | VII | VIII |
|---|---|---|---|---|---|---|---|---|
| "Nouns" | 110 | Language | 59 | 64. | 57.9 | 73.4 | 82.2 | |
| "Abbreviations" | 1200 | Language | 42 | | | | | |
| Immediate recall | | | | | 65.3 | 64.7 | 23.7 | |
| Delayed recall | | | | | 72.7 | 70.5 | 59.7 | |
| "Peanuts" | 1900 | Agriculture | 156 | | | | | |
| Previous knowledge | | | | 10.5 | 15.9 | 21.9 | | |
| Immediate recall | | | | 24.5 | 36.2 | 47.5 | | |
| "Gov't of Switzerland" | 1200 | Civics | 16 | | | | | |
| Previous knowledge | | | | | | 10. | | |
| Immediate recall | | | | | | 43.9 | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Delayed recall | | | | | | 25. | | |
| ‘‘Medieval Castles’’ | 687 | History | 16 | | | | | |
| Previous knowledge | | | | | | 16.3 | | |
| Immediate recall | | | | | | 57.3 | | |
| Delayed recall | | | | | | 27.3 | | |
| ‘‘Rainbow’’ | 1000 | Narrative | 16 | | | | | |
| Previous knowledge | | | | | | 34.8 | | |
| Immediate recall | | | | | | 78. | | |
| Delayed recall | | | | | | 70.3 | | |
| ‘‘Tuberculosis’’ | 1900 | Hygiene | 16 | | | | | |
| Previous knowledge | | | | | | 14.8 | | |
| Immediate recall | | | | | | 31.6 | | |
| Delayed recall | | | | | | 21.1 | | |
| ‘‘Admiralty Islanders’’ | 1000 | Geography | 475 | | | | | |
| Previous knowledge | | | | 6. | 11.2 | 12.3 | 18.2 | 19.3 |
| Immediate recall | | | | 15.5 | 25. | 29.5 | 37.7 | 45.3 |
| Delay recall | | | | 13.5 | 27. | 28. | 32.2 | 38.6 |
| ‘‘Abbreviations’’ | 1200 | Language | 26 | | | | | |
| Previous knowledge | | | | | | | | 28. |
| Immediate recall | | | | | | | | 31. |
| Delayed recall | | | | | | | | 45. |
| ‘‘Medieval Castles’’ | 687 | History | 228 | | | | | |
| Immediate recall | Group A | | | | | | 24. | 37. |
| Delayed recall | Group A | | | | | | 13.7 | 23.9 |
| Immediate recall | Group B | | | | | | 27.5 | 38.1 |
| Delayed recall | Group B | | | | | | 15.2 | 19.7 |

In the foregoing table the scores are expressed as per cents of the total possible score. They are here presented in convenient form for the purpose of comparative study and in order that the reader may get a summarized view of the present study as a whole. It will be seen that the table contains the following:

1) Experiments involving materials of different character and length; *viz.*, language, agriculture, civics, history, narrative, hygiene, geography.

2) Materials varying in length from 110 words to 1900 words.

3) Number of pupils examined at one time varying from 16 to 475.

4) Sixteen pupils used in five different experiments.

5) Significant experiments involving a considerable number of children, made with agriculture, history, and geographical materials.

6) Scores varying significantly with the types of materials used.

7) In six of these experiments, pupils tested three times; once for previous knowledge, once for immediate recall and once for delayed recall.

8) In the other experiments, at least two tests given.

Comparisons of these data with previous studies will follow in Table 45.

TABLE 45

COMPARISON OF PRESENT STUDY WITH PREVIOUS STUDIES

| Name of Study | Grades IV | V | VI | VII | VIII | Character of Material |
|---|---|---|---|---|---|---|
| Pintner & Gilliland | 64. | | | | | Short Par. |
| Present Study | | | | | | |
| "Nouns" | 64 | | | | | Short Par. |
| Waldo | 60. | 25.6 | 35. | 48.3 | 63.5 | Current Ev. |
| Present Study | | | | | | |
| "Abbreviations" | | | 72.7 | 70.5 | 79.5 | Language |
| "Admiralty Island." | 15.5 | 25. | 29.5 | 37.7 | 45.3 | Geography |
| "Peanuts" | 24.5 | 36.2 | 47.5 | | | Agriculture |
| Brown | | | 27.6 | | | Simple Prose |
| Present Study | | | | | | |
| "Rainbow" | | | 78. | | | Narrative |
| "Peanuts" | | | 47.5 | | | Agriculture |
| "Tuberculosis" | | | 31.6 | | | Hygiene |
| "Medieval Castles. | | | 37.3 | | | History |
| "Gov't Switzerland" | | | 43.9 | | | Civics |
| Mead | 40.4 | | | | | Stories |
| Present Study | 24.5 | | | | | Agriculture |
| "Peanuts" | 24.5 | | | | | Agriculture |
| Mead | 23.9 | 43.1 | 44.3 | 38. | 55. | Narrative |
| Present Study | | | | | | |
| "Admiralty Island." | 15.5 | 25. | 29.5 | 37.7 | 45.3 | Geography |
| Peters | 73. | 78. | 58. | | | Narrative |
| Present Study | | | | | | |
| "Rainbow" | | | 78. | | | Narrative |
| Germane | | | | | | |
| "Peanuts" | | | 21.2 | 27.3 | 42.5 | Agriculture |
| Present Study | | | | | | |
| "Peanuts" | 24.5 | 36.2 | 47.5 | | | Agriculture |

From the preceding table it is seen that the present study agrees with previous studies roughly as follows:

1) On short paragraphs with Pintner and Gilliland very closely; present study 64, Pintner and Gilliland 68, in grade 4.

2) With Waldo in Grades 4 to 6. The present study is slightly higher on language than Waldo's pupils on "Current Events." On geography the present study shows lower scores than Waldo on "Current Events."

3) Brown, in sixth grade, lower on simple prose than present study on narrative, science, agriculture, history, and civics material. This is probably due to the material and to the difference in the tests used.

4) Mead, on stories, about twice the scores shown by the present study on agriculture, in the fourth grade. This is

probably due to the differences in the materials as well as to differences in the tests.

5) Peters with sixth grade pupils, about seventy-five per cent as high as the present study. This is probably due to the type of test used.

6) On factual material, Germane's score in grades 6, 7, and 8, lower in proportion than those in the present study with grades 4, 5, and 6. This difference is attributed to the type of test used and also to the fact that in the present study an initial test was given for previous knowledge which directed the reading. When corrected for the effect of the test the scores agree closely.

The conclusions drawn from the foregoing comparisons are as follows:

1) There is a general agreement among all these studies that scores on narrative materials run high.

2) Scores on factual materials run lower than scores on narrative materials.

3) When allowance is made for the effect of giving a test before reading, the scores in the present study run lower than those in previous studies.

4) The tests used in the present study are designed to make less demand upon power of unstimulated recall and composition ability and thus tend to produce higher scores.

5) Tentative norms for geographical and agricultural material, History and short paragraphs are established by the present this.

6) There was some material bearing on the present problem available in previous studies but it had not been gathered with the purpose of measuring the effect of a single reading.

7) The effect of the above summary is to bring out the importance of a single reading and to point out the contributions given by the scattered data of other investigators.

8) The present study furnished data on two readings. These data are given by no previous investigator discoverable by the present experimenter.

## GENERAL CONCLUSIONS FROM THE PRESENT STUDY

The following conclusions stand out as the most important general principles deduced from the present study or clearly indicated by the results of the data presented:

1) The effect of a single reading of content material varies with individuals but on the average it is less than half of the total ideas in the article read and it often falls to a third or less.

2) The effect of a single reading as measured by delayed recall shows that while an indispensable part of the process of reading and remembering, it is much less powerful as a means of insuring memory than the review incident to an immediate recall test.

3) The effects of a single reading vary with the type of test used. Almost any kind of result might be obtained by varying the type of test and the system of scoring the answers to the test.

4) If a single reading takes place without any immediate recall of the ideas, it is doubtful that the reading will leave any impression in the mind of the learner after a lapse of twenty or thirty days, unless the material is very interesting or striking.

5) The effect of a single reading varies in efficiency with the grade taking the test. Pupils in the lower grades show much less ability than those in the upper grades; but there are individuals in all the grades who overlap considerably the grades above them in their ability to perform the tests after a single reading.

6) The consistent advantage of giving an initial test to determine the previous knowledge of the pupil suggests that the preliminary test before reading may be an excellent method of teaching pupils to read efficiently because it directs the reading.

7) If learners were limited to a single contact with reading material, most individuals would probably be very inefficient, or would be forced to develop a very high degree of skill in this process. Nor does a second reading of the same type as the first add materially to the knowledge of the learner. Since the effects of repetition and recall are so well brought out by these experiments, the implication is that after several months the effect of a single reading of factual material of this sort would hardly be noticeable at all.

8. Individuals vary so much in their abilities that it is not safe to conclude that there are not some people who can read an article once and practically remember its entire contents. There are a fortunate few who would be as efficient after a

single reading as others might be after severe study. Macauley is an example of the former type.

9) This study emphasizes the need for the determination of the relative difficulty of different types of materials under varying conditions, and for devising methods to secure better retention than can be possible after a single hasty reading.

## SELECTED BIBLIOGRAPHY

1. Huey, E. B., *Psychology and Pedagogy of Reading.* Macmillan Company, 1908.
2. Courtis, S. A., *Measurement of Classroom Products,* Gary Survey, pp. 263-265, General Educational Board, 1919.
3. Watkins, Emma, *How to Teach Silent Reading to Beginners.* Lippincott, 1922.
4. Stone, C. R., *Silent and Oral Reading.* Houghton Mifflin Company, 1922.
5. Ebbinghaus, E., *"Uber Das Gedächtnis"* (Translated by Ruger), Teachers College, Columbia University, 1912.
6. Meumann, E., *The Psychology of Learning* (Translated by John Wallace Baird). D. Appleton, 1913.
7. Thorndike, E. L., *Eduactional Psychology,* Vol. II, Teachers College, Columbia University, 1913.
8. Edwards, A. S., *Fundamental Principles of Learning and Study.* Warwick and York, Inc., 1920.
9. Whipple, G. M., *Manual of Mental and Physical Tests,* Part II, Warwick and York, Inc., 1915.
10. Pintner, Rudolph, "Oral and Silent Reading", *Journal of Educational Psychology,* IV: 333-337.
11. Whipple and Courtis, "Preliminary Investigation of Skimming in Reading", *Journal of Educational Psychology,* VII: 333-349.
12. Waldo, K. D., "Tests of Reading in the Sycamore Schools,", *Elementary School Journal,* XV: 231ff.
13. Brown, H. A., "Efficiency in Instruction in Reading", *Elementary School Journal,* VII: 201-212.
14. Gray, W. S., *Studies of Elementary School Reading through Standardized Tests,* Supplementary Educational Monographs, Vol. I, No. 1, University of Chicago Press.
18. Pintner and Gilliland, "Oral and Silent Reading", *Journal of Educational Psychology,* VII: 201-212.
19. Mead, C. D., "Silent versus Oral Reading with 100 Sixth Grade Children", *Journal of Educational Psychology,* Vol. VI.
20. Peters, Chas. C., "Influence of Speed Drills on Rate and Effectiveness of Silent Reading", *Journal of Educational Psychology,* VIII: 350-366.
21. Germane, Edith G., *The Correlation between Speed and Comprehension in Silent Reading.* M. A. Thesis, State University of Iowa, 1920. Unpublished.
22. Courtis, S. A., *Measurement of Classroom Products.* Gary Survey, pp. 274, 315.
23. Germane, Chas. E., *The Value of Summarizing as Compared with Rereading the same Article.* Ph. D. Thesis, Iowa State University, 1920. Chap. II.
25. Alderman, G. H., *The Lecture Method of Teaching versus the Question and Answer Method.* Ph. D. Thesis, State University of Iowa, 1920. Unpublished.

26. Pressey, L. W. & S. L., "A Critical Study of the Concept of Silent Reading", *Journal of Educational Psychology*, XII: 25-31.
27. Rugg, H. O., *Statistical Methods Applied to Education.* Houghton Mifflin Company, 1917.
28. Hoover, J. H., "Motivated Drill Work in Third Grade Silent Reading", *Twentieth Year Book of the National Society for the Study of Education*, Part II, pp. 77-89.
29. Yoakam, G. A., "The Effect of a Single Reading", *Twentieth Year Book of the National Society for the Study of Education*, Part II, pp. 90-102, 1921.
30. Burgess, May Ayres, "Measurement of Silent Reading Ability", *Twentieth Year Book of the National Society for the Study of Education*, Part II, 25-38.
31. Pressey and Pressey, "The Relative Value of Rate and Comprehension Scores of Monroe's Silent Reading Tests as a Measure of Reading Ability", *School and Society*, II:741-748.
32. Germane and Germane, *Silent Reading.* Row Peterson and Company, 1922.

S.C.E.R.T., West Bengal
Library
Calcutta
25/3, B. C. Road

Obtainable from the University
Editor Price, $0.75